EXHIBIT 1

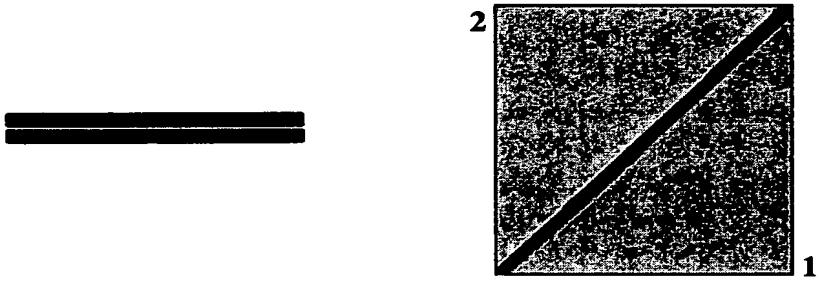
NCBI **Blast 2 Sequences results**

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 34862435	similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) [Rattus norvegicus]	**Length** 804 (1 .. 804)
Sequence 2	gi 15030324	Tumor rejection antigen gp96 [Mus musculus]	**Length** 802 (1 .. 802)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1550 bits (4014), Expect = 0.0
Identities = 788/805 (97%), Positives = 796/805 (97%), Gaps = 4/805 (0%)
```

```
Query: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALAGNEELTVKIKCDREKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENALAGNEELTVKIKCD+EKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALAGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDETAQ 300
           LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDE A+
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDEAAK 300

Query: 301 EEKEEADDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEE+DDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
```

```
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKSKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEK+KESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRVKEDEDDKTVMDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRR+KEDEDDKTVMDLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVMDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTTEDTTDDSEQDE-E 779
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDT+E+  +DSEQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTSEE-AEDSEQDEGE 779

Query: 780 ETDAGAEEEEEEQETEKEPTEKDEL 804
           E DAG EEEEE  ETEKE TEKDEL
Sbjct: 780 EMDAGTEEEEE--ETEKESTEKDEL 802

CPU time:     0.05 user secs.       0.00 sys. secs          0.05 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7160
Number of extensions: 3828
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 804
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 666
Effective length of database: 666,719,727
```

```
Effective search space: 444035338182
Effective search space used: 444035338182
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

Blast 2 Sequences results

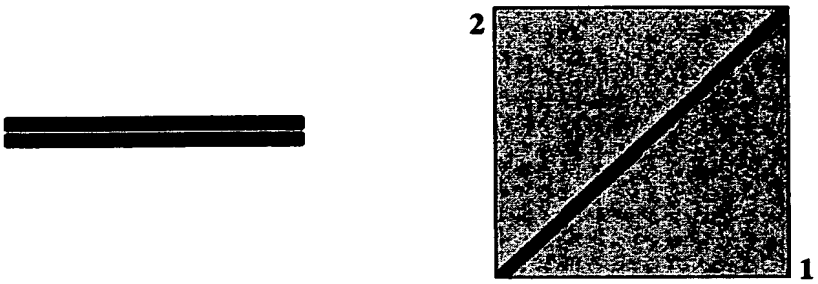
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1 gi 44890631 Tumor rejection antigen (gp96) 1 [Homo sapiens] **Length** 803 (1 .. 803)

Sequence 2 gi 15030324 Tumor rejection antigen gp96 [Mus musculus] **Length** 802 (1 .. 802)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1538 bits (3983), Expect = 0.0
Identities = 775/803 (96%), Positives = 791/803 (97%), Gaps = 1/803 (0%)
```

```
Query: 1   MRALWVLGLCCVLLTFGSVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MR LWVLGLCCVLLTFG VRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALSGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENAL+GNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALAGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTETVEEPMEEEEAAK 300
           LGRGTTITLVLKEEASDYLELDTIKNLV+KYSQFINFPIYVWSSKTETVEEP+EE+EAAK
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDEAAK 300

Query: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
```

```
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADDKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIAD+KY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHH TDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEKTKESREA EKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTV+DLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVMDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDEDE 780
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDP+A+VEEEPEEEPE+T+E+  ED+EQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTSEE-AEDSEQDEGE 779

Query: 781 EMDVGTDEEEETAKESTAEKDEL 803
           EMD GT+EEEE  ++ + + EKDEL
Sbjct: 780 EMDAGTEEEEEETEKESTEKDEL 802

CPU time:     0.05 user secs.       0.01 sys. secs          0.06 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7092
Number of extensions: 3817
Number of successful extensions: 61
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
```

```
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

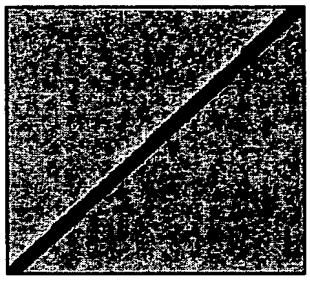
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 34862435	similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) [Rattus norvegicus]	**Length** 804 (1 .. 804)
Sequence 2	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)



RAT 96

HUMAN 96

NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 1527 bits (3954), Expect = 0.0
Identities = 773/806 (95%), Positives = 790/806 (97%), Gaps = 5/806 (0%)

```
Query: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MR LWVLGLCCVLLTFG VRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRALWVLGLCCVLLTFGSVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALAGNEELTVKIKCDREKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENAL+GNEELTVKIKCD+EKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALSGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDETAQ 300
           LGRGTTITLVLKEEASDYLELDTIKNLV+KYSQFINFPIYVWSSKTETVEEP+EE+E A+
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTETVEEPMEEEEAAK 300

Query: 301 EEKEEADDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEE+DDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
```

```
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIAD+KY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADDKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHH TDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKSKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEK+KESREA EKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRVKEDEDDKTVMDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRR+KEDEDDKTV+DLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTTEDTTDDSEQDE-E 779
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDP+A+VEEEPEEEPE+T EDTT+D+EQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDEDE 780

Query: 780 ETDAGAEEEEEEQETEKEPT-EKDEL 804
           E D G +EEE   ET KE T EKDEL
Sbjct: 781 EMDVGTDEEE---ETAKESTAEKDEL 803

CPU time:     0.05 user secs.       0.01 sys. secs          0.06 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7146
Number of extensions: 3810
Number of successful extensions: 63
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 804
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 666
Effective length of database: 666,719,727
```

```
Effective search space: 444035338182
Effective search space used: 444035338182
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

EXHIBIT 2

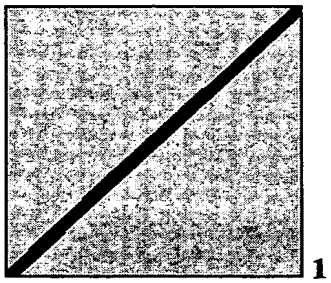
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1436 bits (3717), Expect = 0.0
Identities = 725/733 (98%), Positives = 731/733 (98%), Gaps = 1/733 (0%)
```

```
Query:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
              MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
Conflict  7         *
modified  7         *
modified  5       *
HSPCA     1   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    61  YETLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
              YE+LTDPSKLDSGKELHINLIP+KQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct:    61  YESLTDPSKLDSGKELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
HSPCA     61  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
              ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct:    121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
HSPCA     121 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKED 240
              GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKE+
Sbjct:    181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
modified  231                                          *
HSPCA     181 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
```

```
Query:    241 KEEEKEKEEKESEDKPEIEDVGSDEEEE-KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 299
              KEEEKEKEEKES+DKPEIEDVGSDEEEE KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
modified  263                          *
Conflict  243   ****
HSPCA     241 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    300 NPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 359
              NPDDITNEEYGEFYKSLTNDWE+HLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
Conflict  356                                                          *
HSPCA     301 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    360 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 419
              NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
HSPCA     361 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    420 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 479
              CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
HSPCA     421 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    480 YCTRMKENQKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 539
              YCTRMKENQKHIY+ITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
HSPCA     481 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    540 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 599
              TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
HSPCA     541 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    600 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 659
              VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
HSPCA     601 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    660 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPL 719
              KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPT DDTSAAVTEEMPPL
Sbjct:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
HSPCA     661 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    720 EGDDDTSRMEEVD 732
              EGDDDTSRMEEVD
Sbjct:    721 EGDDDTSRMEEVD 733
HSPCA     721 +++++++++++++

CPU time:     0.04 user secs.       0.00 sys. secs          0.04 total secs.

Lambda      K       H
   0.312    0.131    0.361

Gapped
Lambda      K       H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

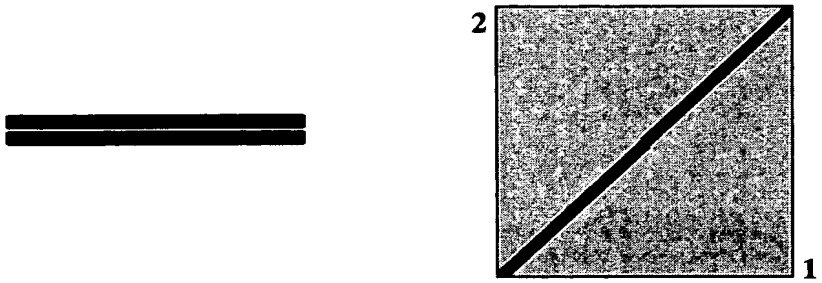
```
Number of Sequences: 1
Number of Hits to DB: 5682
Number of extensions: 3005
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 48734827	Heat shock protein 1, alpha [Rattus norvegicus]	**Length** 733 (1 .. 733)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1450 bits (3754), Expect = 0.0
Identities = 732/733 (99%), Positives = 733/733 (99%)
```

```
Query:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
              MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
Conflict  7         *
modified  7         *
modified  5       *
HSPCA     1   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   61   YESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
              YESLTDPSKLDSGKELHINLIP+KQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct:   61   YESLTDPSKLDSGKELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
HSPCA    61   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   121  ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
              ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct:   121  ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
HSPCA    121  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   181  GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
              GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE
Sbjct:   181  GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
modified 231                                                  *
HSPCA    181  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
```

```
Query:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
              KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
modified 263                      *
Conflict 243   ****
HSPCA     241 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
              NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
Conflict 356                                                          *
HSPCA     301 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
              NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
HSPCA     361 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
              CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
HSPCA     421 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
              YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
HSPCA     481 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
              TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
HSPCA     541 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
              VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
HSPCA     601 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
              KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL
Sbjct:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
HSPCA     661 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    721 EGDDDTSRMEEVD 733
              EGDDDTSRMEEVD
Sbjct:    721 EGDDDTSRMEEVD 733
HSPCA     721 +++++++++++++

CPU time:      0.04 user secs.       0.01 sys. secs          0.05 total secs.

Lambda      K       H
   0.312    0.131    0.360

Gapped
Lambda      K       H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

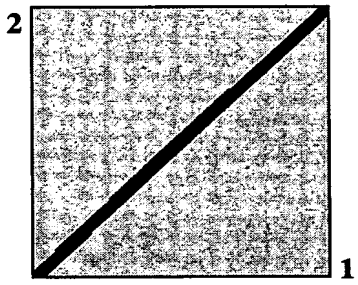
```
Number of Sequences: 1
Number of Hits to DB: 5739
Number of extensions: 3027
Number of successful extensions: 71
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 733
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 596
Effective length of database: 666,719,728
Effective search space: 397364957888
Effective search space used: 397364957888
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 48734827	Heat shock protein 1, alpha [Rattus norvegicus]	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1438 bits (3722), Expect = 0.0
Identities = 726/733 (99%), Positives = 731/733 (99%), Gaps = 1/733 (0%)
```

```
Query: 1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
           MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct: 1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60

Query: 61  YETLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
           YE+LTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct: 61  YESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120

Query: 121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
           ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct: 121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180

Query: 181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKED 240
           GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKE+
Sbjct: 181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240

Query: 241 KEEEKEKEEKESEDKPEIEDVGSDEEEE-KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 299
           KEEEKEKEEKES+DKPEIEDVGSDEEEE KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct: 241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300

Query: 300 NPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 359
           NPDDITNEEYGEFYKSLTNDWE+HLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct: 301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
```

```
Query: 360 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 419
           NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct: 361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420

Query: 420 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 479
           CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct: 421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480

Query: 480 YCTRMKENQKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 539
           YCTRMKENQKHIY+ITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct: 481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540

Query: 540 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 599
           TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct: 541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600

Query: 600 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 659
           VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct: 601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660

Query: 660 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPL 719
           KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPT DDTSAAVTEEMPPL
Sbjct: 661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720

Query: 720 EGDDDTSRMEEVD 732
           EGDDDTSRMEEVD
Sbjct: 721 EGDDDTSRMEEVD 733

CPU time:     0.06 user secs.       0.00 sys. secs          0.06 total secs.

Lambda     K      H
   0.312    0.131    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5690
Number of extensions: 3010
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
```

:3: 129 (50.0 bits)
:1: 42 (22.0 bits)
:2: 80 (35.4 bits)

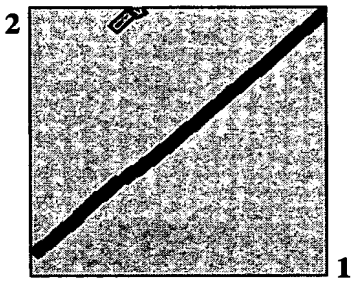
EXHIBIT 3

Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 638 bits (1645), Expect = 0.0
Identities = 344/729 (47%), Positives = 479/729 (65%), Gaps = 27/729 (3%)

```
Query: 15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYETLTDPSKLDSGK 74
           E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  +LTD + L   +
Sbjct: 71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130

Query: 75  ELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129
           EL + +  +K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct: 131 ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190

Query: 130 MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189
           +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct: 191 LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250

Query: 190 LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEEKEKEE 249
           LKE+ ++YLE   IK +VKK+SQFI +PI ++  K          E  E+  +EEE  KEE
Sbjct: 251 LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT--------ETVEEPMEEEEAAKEE 302

Query: 250 KESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEEY 309
           KE  D    ++   +EEEE+K   K K KK+++   D E +N  KPIW R   ++  +EY
Sbjct: 303 KEESD----DEAAVEEEEEEK---KPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEY 355

Query: 310 GEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVRR 367
             FYKS + + +D +A  HF+ EG++ F+++LFVP  AP  LF+    KK + IKLYVRR
Sbjct: 356 KAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRR 415
```

```
Query: 368 VFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTEL 427
           VFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  ++
Sbjct: 416 VFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKI 475

Query: 428 AEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKEN 487
           A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct: 476 ADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEK 534

Query: 488 QKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTKE 547
           Q  IY++ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V KE
Sbjct: 535 QDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKE 594

Query: 548 GLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYGW 606
           G++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YGW
Sbjct: 595 GVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGW 654

Query: 607 TANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDLV 663
           + NMERIMKAQA    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct: 655 SGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLA 714

Query: 664 ILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPLEGDD 723
           ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++       E        D
Sbjct: 715 VVLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDT 774

Query: 724 DTSRMEEVD 732
           +    EE+D
Sbjct: 775 EQDEDEEMD 783

Score = 35.4 bits (80), Expect = 8.6
Identities = 21/73 (28%), Positives = 38/73 (51%), Gaps = 2/73 (2%)

Query: 208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEE 267
           K +   I   + L +  + D +V ++  EE E+  E+  ++ ++ ED  E  DVG+DEEE
Sbjct: 733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790

Query: 268 EKKDGDKKKKKKI 280
           E       +K ++
Sbjct: 791 ETAKESTAEKDEL 803

CPU time:     0.08 user secs.       0.00 sys. secs          0.08 total secs.

Lambda     K      H
   0.312    0.131    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5882
Number of extensions: 3409
Number of successful extensions: 66
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 5
```

```
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

Blast 2 Sequences results

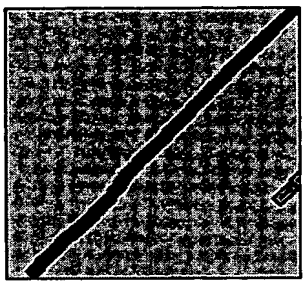
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

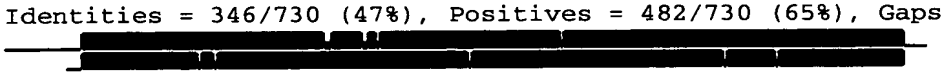
Sequence 1 gi 44890631 Tumor rejection antigen (gp96) 1 [Homo sapiens] **Length** 803 (1 .. 803)

Sequence 2 gi 48734827 Heat shock protein 1, alpha [Rattus norvegicus] **Length** 733 (1 .. 733)



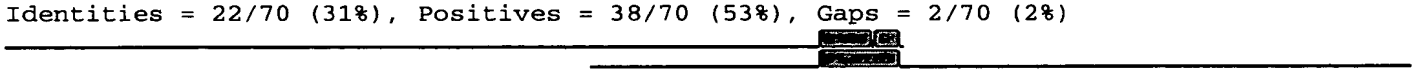
NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 641 bits (1654), Expect = 0.0
Identities = 346/730 (47%), Positives = 482/730 (65%), Gaps = 28/730 (3%)



```
Query: 71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130
           E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  SLTD + L   +
Sbjct: 15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGK 74

Query: 131 ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190
           EL + +  +K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct: 75  ELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129

Query: 191 LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250
           +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct: 130 MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189

Query: 251 LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT-------ETVEEPMEEEEAAKEEK 303
           LKE+ ++YLE   IK +VKK+SQFI +PI ++  K         +  EE  E+EE  ++E+
Sbjct: 190 LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEE 249

Query: 304 EESDDEAAVE-----EEEEEKKP----KTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDE 354
           +ESDD+  +E     EEEEEKK     K KK+++   D E +N  KPIW R   ++  +E
Sbjct: 250 KESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEE 309

Query: 355 YKAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVR 414
           Y  FYKS + + ++ +A  HF+ EG++ F+++LFVP  AP  LF+    KK + IKLYVR
Sbjct: 310 YGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVR 367

Query: 415 RVFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKK 474
           RVFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  +
Sbjct: 368 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFSE 427
```

```
Sbjct: 368 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTE 427

Query: 475 IADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKE 533
           +A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct: 428 LAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKE 487

Query: 534 KQDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAK 593
            Q  IYF+ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V K
Sbjct: 488 NQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTK 547

Query: 594 EGVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYG 653
           EG++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YG
Sbjct: 548 EGLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYG 606

Query: 654 WSGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDL 713
           W+ NMERIMKAQA    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct: 607 WTANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDL 663

Query: 714 AVVLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTED 773
            ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++       E       D
Sbjct: 664 VILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPLEGD 723

Query: 774 TEQDEDEEMD 783
            +    EE+D
Sbjct: 724 DDTSRMEEVD 733

Score = 38.5 bits (88), Expect = 1.1
Identities = 22/70 (31%), Positives = 38/70 (53%), Gaps = 2/70 (2%)
```



```
Query: 733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790
           K +   I   + L +  + D +V ++  EE EE  E+  ++ ++ +D  E  DVG+DEEE
Sbjct: 208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEEKESDDKPEIEDVGSDEEE 267

Query: 791 ETAKESTAEK 800
           E  K+   +K
Sbjct: 268 EEKKDGDKKK 277

CPU time:     0.11 user secs.       0.01 sys. secs          0.12 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5960
Number of extensions: 3415
Number of successful extensions: 68
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 16
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
```

```
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

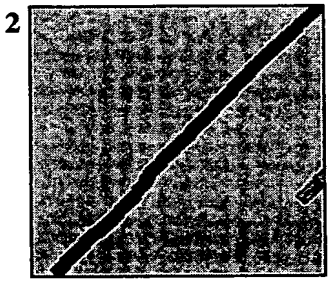
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

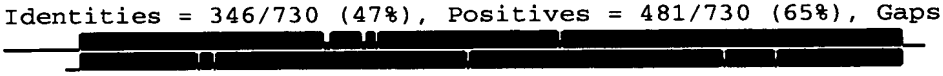
Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



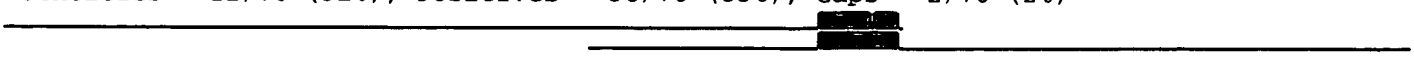
NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score =  641 bits (1653), Expect = 0.0
Identities = 346/730 (47%), Positives = 481/730 (65%), Gaps = 28/730 (3%)
```



```
Query:    71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130
              E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  SLTD + L   +
Sbjct:    15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGK 74
HSPCA     15  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    131 ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190
              EL + +   K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct:    75  ELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129
HSPCA     75  ++++++++++++++++++++++++++++++++++++++++++++    ++++++++++

Query:    191 LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250
              +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct:    130 MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189
HSPCA     130 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    251 LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT-------ETVEEPMEEEEAAKEEK 303
              LKE+ ++YLE   IK +VKK+SQFI +PI ++  K         +  EE  E+EE  ++E+
Sbjct:    190 LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEE 249
HSPCA     190 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified  231                                    *
Conflict  243                                                  ****

Query:    304 EESDDEAAVE-----EEEEEKKP----KTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDE 354
              +ESDD+  +E     EEEEEKK     K KK+++   D E +N  KPIW R   ++  +E
```

```
Sbjct:    250 KESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEE 309
HSPCA     250 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified  263              *

Query:    355 YKAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVR 414
              Y  FYKS + + ++ +A  HF+ EG++ F+++LFVP  AP  LF+    KK + IKLYVR
Sbjct:    310 YGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVR 367
HSPCA     310 ++++++++++++++++++++++++++++++++++++++++++++++  ++++++++++++
Conflict  356                                                 *

Query:    415 RVFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKK 474
              RVFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  +
Sbjct:    368 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTE 427
HSPCA     368 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    475 IADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKE 533
              +A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct:    428 LAEDKENYKKFYEQFSKNIKLGIIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKE 487
HSPCA     428 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    534 KQDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAK 593
               Q  IYF+ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V K
Sbjct:    488 NQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTK 547
HSPCA     488 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    594 EGVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYG 653
              EG++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YG
Sbjct:    548 EGLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYG 606
HSPCA     548 ++++++++++++++++++++++++++++++ ++++++++++++++++++++++++++++

Query:    654 WSGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDL 713
              W+ NMERIMKAQA.    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct:    607 WTANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDL 663
HSPCA     607 ++++++++++++++   ++++++++++++++++++++++++++++++++++++++++++++

Query:    714 AVVLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTED 773
               ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++      E       D
Sbjct:    664 VILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPLEGD 723
HSPCA     664 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    774 TEQDEDEEMD 783
               +    EE+D
Sbjct:    724 DDTSRMEEVD 733
HSPCA     724 ++++++++++

Score = 38.5 bits (88), Expect = 1.1
Identities = 22/70 (31%), Positives = 38/70 (53%), Gaps = 2/70 (2%)
```



```
Query:    733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790
              K +   I   + L +  + D +V ++  EE EE  E+  ++ ++ +D  E  DVG+DEEE
Sbjct:    208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEEKESDDKPEIEDVGSDEEE 267
Conflict  243                                    ****
modified  231                        *
HSPCA     208 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified  263                                                      *

Query:    791 ETAKESTAEK 800
              E  K+   +K
```

```
Sbjct:   268 EEKKDGDKKK 277
HSPCA    268 ++++++++++

CPU time:     0.10 user secs.       0.00 sys. secs          0.10 total secs.

Lambda      K       H
   0.312    0.130    0.361

Gapped
Lambda      K       H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5958
Number of extensions: 3413
Number of successful extensions: 68
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 16
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

EXHIBIT I

Annu. Rev. Genet. 1988. 22:631-77

THE HEAT-SHOCK PROTEINS

S. Lindquist

Department of Molecular Genetics and Cell Biology, The University of Chicago, Chicago, Illinois 60637

E. A. Craig

Department of Physiological Chemistry, The University of Wisconsin, Madison, Wisconsin 53706

CONTENTS

0066-4197/88/1215-0631$02.00

PERSPECTIVES

All organisms respond to heat by inducing the synthesis of a group of proteins called the heat-shock proteins or *hsps*. The response is the most highly conserved genetic system known, existing in every organism in which it has been sought, from archaebacteria to eubacteria, from plants to animals. Although certain features of the response vary from organism to organism, many are universal, or nearly so. All organisms examined produce proteins encoded by the hsp70 and hsp90 gene families in response to elevated temperatures. These proteins are among the most highly conserved proteins in existence. Also universally, several of the proteins induced by heat are induced by a variety of other stresses. Although the particular constellation of effective inducers varies somewhat from organism to organism, in nearly all cells anoxia, ethanol, and certain heavy metal ions induce the proteins. Furthermore, either the hsps themselves or their close relatives are present in all organisms at normal temperatures and play vital roles in normal cell function. This last finding has provided important information on the specific molecular functions of the proteins and will be discussed in detail in this review.

An early and long-standing assumption about the heat-shock response is that the hsps protect cells from the toxic effects of heat and other stresses; good evidence supports this view. There is, first of all, the very nature of the response. In all organisms, the induction of hsps is remarkably rapid and intense, in keeping with the notion that it is an emergency response. Moreover, there is a striking relationship between the induction temperature and the organism's environment. In different organisms the response is induced at very different temperatures. In each case, the organism would be expected to cope with such temperatures in its natural environment. Thus, in the fruit fly *Drosophila melanogaster*, induction occurs between 33–37°C, common temperatures on warm summer days (125). In thermophilic bacteria growing at 50°C, the proteins are induced when temperatures are raised to 60°C (51). In arctic fishes growing at 0°C, they are induced at 5–10°C (B. Maresca, personal communication). In mammals they are induced by fever temperatures (122), and in soybeans they are induced in the field on hot sunny days (104).

A particularly interesting example is provided by a variety of dimorphic pathogens that cycle between relatively cool temperatues in one phase of their life cycle and the warmer temperatures of their mammalian hosts in another phase. This change in temperature is accompanied by the strong induction of hsps, in both prokaryotic and eukaryotic pathogens. In what is almost certainly a related phenomenon, heat-shock proteins are immunodominant antigens in many of these infections. Both circulating antibodies and activated T cells have been shown to have specificity for the major heat-shock proteins

of organisms as diverse as *Mycobacterium leprae* (the causative agent in leprosy), *M. tuberculosis* (tuberculosis), *Coxiella burnetti* (Q fever), *Plasmodium falciparum* (malaria), *Schistosoma mansoni* (schistosomiasis), *Brugia malayii* (filariasis) (Selkirk et al), *Trypanosoma cruzi* (Chagas disease) and *Leishmania major* (Leishmaniasis) (reviewed in 247). In the case of *P. falciparum*, hsp70 was reported to be on the cell surface, but this has been disputed (6, 16). Given the fact that so many different hsps, not generally on the cell surface, have been reported as immunodominant antigens, the most likely explanation for their antigenicity is that they are extremely abundant proteins at high temperatures and are therefore processed by macrophages as major foreign antigens for presentation to lymphocytes. The induction of the heat-shock proteins might be further enhanced by the hostile environment in the macrophages which engulf them and which serve as host cells for some of the organisms.

The most compelling argument that hsps have a protective function is the strong correlation between their induction and the induction of thermotolerance. Thermotolerance experiments with similar design and general result have been performed in both cultured cells and a wide variety of organisms, including bacteria, yeasts, slime molds, soybean seedlings, fruit flies, and mice. The basic observation is that a group of cells or organisms is killed rapidly when shifted directly from its normal growing temperature to a much higher temperature, whereas a matched group, given a mild preheat treatment to induce hsps, is killed much more slowly. Moreover, preheat treatments induce tolerance to other forms of stress, and other forms of stress induce tolerance to heat. Apparently then, the heat shock proteins are induced by moderate stresses, stresses which themselves are not necessarily lethal, in order to protect the organism from even more severe stress. The system seems to make eminent biological sense.

Unfortunately, this simple picture dissolves on closer inspection. There are circumstances in which the induction of thermotolerance does not correlate with the induction of heat-shock proteins, and attempts to identify the role of any one protein in thermotolerance have met widely with frustration. Fortunately, much progress has come from unexpected quarters. The hsps themselves, or close relatives produced at normal temperatures, serve vital functions in normal cells. Their roles in normal cell function are providing important clues to their putative functions in thermotolerance. In this review we first describe what is known about individual proteins, concentrating on a few systems in which biochemical or genetic analyses have been particularly fruitful. We regret that the broad scope of the current heat-shock literature and the limits on the length of this review preclude the consideration of many interesting findings from other systems. We then examine the hypothesis that hsps play a vital role in thermotolerance.

HSP110

Most, but by no means all, eukaryotes produce proteins of greater than 100 kd in response to high temperatures. They have been characterized in detail only in mammalian cells. The 110 kd protein of murine cells is found in the nucleus, with concentration in the nucleolus of both control and heat-shocked cells (204). The protein separates from the phase-dense nucleolar body, forming a nucleolar cap, when cultures growing at normal temperatures become confluent or when actively growing cells are incubated without serum or are treated with actinomycin (195). Brief heat shocks do not lead to nucleolar segmentation in proliferating cells; in confluent cultures they reverse it. With longer heat shocks, hsp110 forms a ring-like structure at the nucleolar periphery (239). Immunoelectron microscopy indicates that hsp110 associates with the fibrillar component of nucleoli, the site of nucleolar chromatin (rDNA). Treatment of fixed cells with RNase eliminates staining (204), suggesting that the protein associates with RNA or with a complex of proteins that bind RNA. Since ribosome production is very sensitive to heat shock (155), it is speculated that hsp110 is induced to protect it. In this respect it is notable that a member of the mammalian hsp70 gene family also localizes to nucleoli and has been postulated to protect ribosome assembly (121, 166, 236). Unlike hsp110, this protein concentrates in the granular region of the nucleolus, the location of pre-ribosomes (238).

Other nucleolar proteins of 110 kd have been studied independently of the heat-shock response (52, 105). The peptide map of one, C23 (195), is similar but not identical to that of hsp110, and this protein, too, has been localized to the fibrillar region of the nucleolus (52). Unfortunately, the gene encoding hsp 110 has not been isolated, and no genetic analysis has been performed. Recently, the 104 kd hsp of *Saccharomyces cerevisiae* has been purified, used to produce antibodies, and shown to be a nuclear protein (K. Borkovich, S. Lindquist, manuscript in preparation). If it should prove to be an analog of the mammalian protein, this gene family will then be open to genetic analysis.

THE HSP90 FAMILY

Members of the *hsp90* gene family have been cloned and sequenced from several evolutionarily diverse organisms, including fruit flies, yeasts, chickens, mammals, trypanosomes, and bacteria. Sequence analysis of the cloned genes demonstrates that the proteins are very highly conserved. The proteins of even the most distantly related eukaryotes have 50% amino-acid identity, and all have greater than 40% identity with the *Escherichia coli* protein (12, 56, 65, 133). In all eukaryotes, a region of extremely high negative-charge density, which itself shows little sequence conservation, is located at the same relative position in the protein. The *E. coli* protein is missing this segment. All of the proteins, including that of *E. coli*, contain another, smaller region of high negative-charge density toward the carboxy terminus. The carboxy-terminal regions of these proteins are generally the most divergent, but the four most-terminal amino acids, *glu-glu-val-asp*, are the same in all eukaryotic hsp90s (in trypanosmes the second *glu* is replaced by *gln*; 56). It is remarkable that this sequence is also found at the carboxy-terminus of the eukaryotic hsp70 proteins. In other respects these proteins have little or no homology. The sequence must serve some important purpose, but what the purpose may be is presently unknown.

In virtually all cells, proteins of the hsp90 family are abundant at normal temperatures and are further induced by heat. In *D. melanogaster*, there appears to be only one gene in this family, *HSP83* (18). In addition to being constitutively expressed and induced by heat and other stresses, *HSP83* is developmentally induced during oogenesis (252). The haploid genome of the budding yeast *S. cerevisiae* has two genes in this family, encoding nearly identical proteins. One, *HSC83*, is constitutively expressed at a high level and is moderately heat-inducible; the other, *HSP83*, is constitutively synthesized at a lower level and is more strongly heat-inducible (K. Borkavich, F. Farrell, D. Finkelstein, S. Lindquist, in preparation). The more heat-inducible yeast protein is also developmentally regulated and accumulates as cells transit into stationary phase or begin to sporulate (110).

In vertebrate cells further diversification of the genes in this family has occurred with at least one encoding a signal sequence to transport the protein across the endoplasmic reticulum (109, 133, 143, 200). This finding may explain an earlier observation that antibodies against this protein stained the golgi (123). As is the case with the proteins of *Drosophila* and yeast, the other members of the vertebrate family appear to be abundant at normal temperatures, soluble, and predominantly cytoplasmic, with some relocalization in nuclei during heat shock (31, 42, 113, 218). A tumor-specific transplantation antigen, Meth A, has recently been identified as hsp90 (215). Although a small portion of the protein is found on the cell surface in this tumor line, it is postulated that it arrives there secondarily, by deposition of protein from lysed cells.

The ER protein is larger than the cytosolic protein, with an apparent Mwt on SDS gels of 94–108 kd versus 87–92 kd for the cytosolic form. It also contains the sequence *glu-glu-val-asp* at the same relative position in the protein, but in this case it is not C-terminal as the sequence extends an additional 24 amino acids (200). Again, the four most-terminal amino acids of the hsp90 ER proteins are identical to those of the hsp70 ER protein, in this case *lys-asp-glu-leu*. Here, the sequence has been shown to provide retention in the ER, preventing secretion, and is shared by other ER proteins (147). The

ER and cytosolic proteins are often not coordinately regulated. The ER protein is induced by glucose starvation (192) and has been named GRP94 (glucose-regulated protein), while the cytosolic version is induced by glucose restoration. The ER proteins are also induced by heat, steroids, and other agents, but their responses vary with cell type (68, 192, 194).

Biochemical Analysis

Biochemical analysis of hsp90, the cytosolic protein, in mammalian and chicken cells indicates that it associates with very different types of proteins but that, remarkably, it may serve a similar function in these different associations. The first protein with which hsp90 was shown to have a specific association was the transforming protein of Rous Sarcoma Virus, $pp60^{v\text{-}src}$ (26, 158). This tyrosine kinase associates with hsp90 and a 50 kd phosphoprotein immediately after it is synthesized. At or about the time it is released from association with hsp90, it is phosphorylated on tyrosine, inserted into the membrane, and activated as a kinase (27, 43). These results led to the proposal that hsp90 binds to the kinase, keeping it soluble and inactive, while it is transported to its proper location in the plasma membrane. (See Figure 1.)

Further evidence supports this hypothesis and suggests it is of more general significance. Five other transforming proteins with tyrosine kinase activity, *yes, fps, fes, fgr*, also form stable complexes with 90 and 50 kd proteins. In some cases this 90 kd protein has been identified as hsp90 (1, 128, 250). The kinases in these complexes are incapable of autophosphorylation (26, 251), a characteristic of the kinase monomer. Moreover, that fraction of the kinase that can be precipitated from cell lysates with anti-hsp90 sera is underphosphorylated. Finally, mutant $pp60^{v\text{-}src}$ proteins that are transformation defective form much more stable complexes with hsp90. If hsp 90 does have a general inactivating or transporting function, it might be expected to associate with the cellular equivalents of these tyrosine kinases, but this has not yet been described.

hsp90 does, however, associate with other cellular kinases. Highly purified preparations of the heme-controlled eIF2-α kinase contain hsp90 as a prominent component (179). In contrast with the tyrosine kinases, hsp90 appears to stimulate the kinase, thereby increasing phosphorylation of eIF2-α and inhibiting protein synthesis in reticulocyte lysates (D. W. Rose, B. Hardesty, W. J. Welch, manuscript submitted). Whether this contributes to translational regulation in heat-shocked cells is unclear. hsp90 also associates with casein kinase II and is a substrate for phosphorylation by that protein in vitro (179). Moreover, highly purified preparations of yeast protein kinase C contain hsp90 (F. O. Fields, J. Thorner, personal communication). Various members of the hsp90 protein family are phosphorylated in vivo in all organisms investigated, by an unknown mechanism of uncertain regulatory significance.

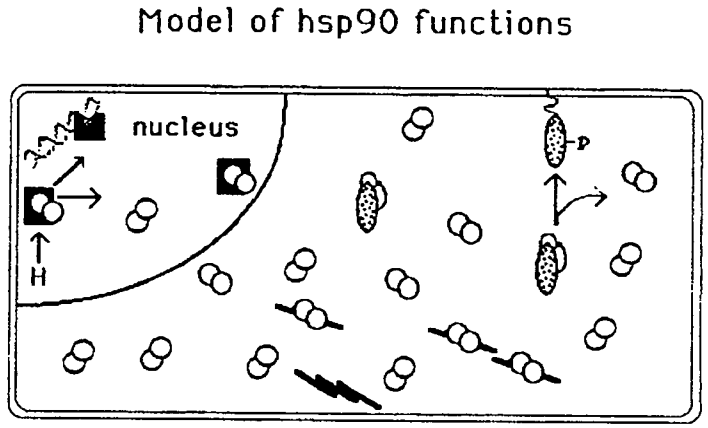


Figure 1 Model of hsp90 functions. The interaction of hsp90 with steroid hormone receptors (dissociated by hormone), tyrosine kinases (prior to insertion in the membrane), and other cellular proteins (including actin and tubulin). Hsp 90 is believed to exist as a dimer and is certainly in excess of the proteins with which it associates. At high temperatures, higher concentrations of hsp90 may be required to maintain the proteins in complexes.

Since hsp90 is not phosphorylated on tyrosine (26, 158), phosphorylation does not appear to be a consequence of its association with tyrosine kinases. That hsp90 is subject to phosphorylation by a DNA-dependent mechanism in vitro is intriguing but, at present, of uncertain import (231).

The other class of hsp90 associations studied in detail is the steroid-hormone receptor complex. All steroid hormone receptors studied to date, including the estrogen, progesterone and glucocorticoid receptors, can be isolated in the unactivated state (that is, in the absence of steroid hormones) as complexes with apparent molecular weights of ~300 kd and sedimentaion coefficients of 8–10S. These rather fragile complexes are dissociated by hormone, high salt (>0.25M KCl), or metal chelators, and are stabilized by low salt and molybdate ions. The complexes contain, in addition to the hormone-binding proteins, 90 kd proteins that have now been identified as hsp90 (33, 55, 98, 174, 175, 185). Since hsp90 is a very abundant protein and exists in vast excess to hormone-binding proteins, their association was initially suspect. However, the universality of the association, its specific stoichiometry (138, 174), and the fact that dissociation of hsp90 from the complex correlates with activation of the receptor for DNA binding (98, 175, 186) provide convincing evidence of its significance.

In the absence of hsp90, the hormone-binding receptor will bind to the DNA whether hormone is present or not (186). hsp90 binds neither DNA nor

hormone. Apparently, binding of hsp90 to the receptor prevents the receptor from binding to DNA until hormone disrupts association of hsp90 to the receptor. The specificity of the dissociated receptor for hormone-responsive DNA elements and its ability to activate transcription have not yet been examined. Although important questions about the role of hsp90 in regulating hormone-receptor activity remain unanswered, it has been postulated that the highly negatively charged domain in the amino-terminal portion of hsp90 forms an α-helix with a charge distribution that resembles the distribution of negatively charged phosphates in DNA. Hormone receptor would then bind to this region of hsp90 instead of binding to DNA, until hormone triggers its release (14, 17, 83). (See Figure 1). In support of this hypothesis, antibodies prepared against this charged domain of hsp90 immunoprecipitate most of the soluble hsp90 from cell extracts, but not the hsp90 that is complexed with steroid-receptor; this suggests that the charged domain is occluded by interaction with receptor (33). For the glucocorticoid receptor, evidence suggests association with hsp90 may inactivate the receptor by keeping it unfolded (256).

It is remarkable that a member of the hsp90 protein family induced by the steroid hormone antheridiol in the fungus *Achlya ambisexualis* is an integral part of the steroid-hormone receptor complex of this organism (28, 196). This suggests that the essential features of the hsp90-receptor interaction will be observed in all steroid-responsive organisms. It is interesting that the putative receptor-binding domain of hsp90 is missing from the hsp90 protein-analog of *E. coli* but is present in proteins of *S. cerevisiae*. There is no evidence that *S. cerevisiae* is naturally responsive to steroid hormones (although an estradiol binding activity has been reported; 29). Thus, if this protein domain should serve the same function in *S. cerevisiae* that it is proposed to serve in vertebrate cells, it may do so by interacting with other transcription factors. Steroid-hormone receptors belong to a large and ancient superfamily of transcription factors that are activated by structurally diverse ligands (63).

In broad outline, hsp90 appears to play a role in steroid receptor complexes similar to that in tyrosine kinase complexes, keeping the receptor inactive until the proper signal for activation is received. The very fact that it interacts with two such different proteins suggests it may interact with others in a similar manner. Recently, hsp90 has been reported to associate with actin in lymphocyte extracts, in a manner that is dependent on calcium and regulated by calmodulin (108, 154). Antibodies directed against hsp90 stain ruffling membranes in these cells, suggesting the interaction also occurs in vivo. It is postulated that the actin association provides a mechanism for transport of hsp90. In this regard, and considering the tendency of hsp90 to move into the nucleus with heat shock, it is intriguing that actin filaments rearrange during heat shock and may even be found in substantial quantities in the nuclei of heat-shocked cells (238). The protein also appears to associate with tubulin

both in vitro and in vivo (24). Given the high concentrations of actin, tubulin, and hsp90 in the cell, it seems likely that these associations are biologically significant.

Genetic Analysis

While biochemical characterization of the hsp90 protein family is more advanced in mammalian and chicken cells, genetic analysis is more advanced in yeast and *E. coli*. In both organisms, cloned genes can be used to construct deletion and disruption mutations in vitro, and these mutations can then be transformed into wild-type cells in a manner that converts the wild-type gene to the mutation. Experiments of this type demonstrate that the *HSP90* gene family is essential in *S. cerevisiae*. Individual mutations in either of the two closely related *HSP83* and *HSC83* genes are viable. Double mutations are lethal. The individual mutants have two interesting phenotypes. First, they do not grow at temperatures above 37°C; 39°C is the maximum growth temperature for the parental strain. (K. Borkovich, F. Farrell, D. Finekistein, S. Lindquist, manuscript in preparation). At 25°C both grow as well as the wild-type. Clearly, the proteins encoded by these two genes serve identical or nearly identical functions, but neither gene alone is adequate for growth at high temperatures. Thus, hsp90 is an essential protein, required in higher concentrations for growth at higher temperatures. This phenotype reflects the pattern of expression observed for this protein family in all organisms: the members of the hsp90 family are abundant, constitutively synthesized proteins that are also heat inducible. The phenotype of yeast mutants defective in *HSP83* and *HSC83* does not tell us whether the protein is required for a different purpose at higher temperatures or is required for the same purpose at all temperatures but in different concentrations. The latter fits more naturally with the proposed function of hsp90 in binding to other proteins and keeping them inactive until they have reached their proper location or until their activity is required. For such interactions to be biologically useful, they must be rather easily disrupted. High temperatures might simply drive the equilibrium more towards dissociation, requiring cells to produce a higher concentration of the protein to achieve the same level of complex formation.

The second phenotype associated with the individual mutations is a reduced ability to withstand exposure to extreme temperatures. They die more rapidly than do wild-type cells during exposure to 50°C. Notably, this difference in thermotolerance is observed in cells that have been grown in acetate, which forces them into respiratory metabolism, but not in cells that have been grown in glucose medium, which supports fermentative metabolism. It may be that hsp90 plays a role in thermotolerance, as classically defined, only in respiring cells. Alternatively, it may play an equally important role in fermenting cells,

but the concentration provided by a single gene may be sufficient for this purpose.

In contrast to the results in yeast, deletion of the *htpG* gene (which encodes the protein denoted C62.5, the only known member of this protein family in *E. coli*) is not lethal (12). However, like the individual deletions in yeast, the mutant does not grow as well as the wild type at high temperatures. The effect of the *htpG* mutation on survival at extreme temperatures has not been tested. The different effects produced by a complete loss of this protein in cells of *S. cerevisiae* and *E. coli* may be explained in several ways. First, another *E. coli* protein might cover the function of C62.5 in the deletion mutant. Arguing against this, no other genes cross-hybridize with the *htpG* gene, even at low stringency; no other proteins cross-react with a polyclonal antibody raised against the protein; and the *htpG* deletion strain does not overproduce any other protein that might thereby be suspected of compensating for its loss. A monoclonal antibody with specificity for mammalian hsp90 has been reported to cross-react with the *E. coli* lon protein (116), but by DNA sequence analysis, the proteins appear unrelated (A. Goldberg, personal communication).

A second possibility is that the protein performs a function that is essential in higher cells but not in bacteria. Simplistically, if the protein serves to ferry other proteins around the cell, as the biochemical evidence strongly suggests, the smaller size of bacterial cells might make this function valuable but not essential. Third, it is possible that the eukaryotic protein has acquired additional functions and that it is these newer functions that are essential. Again, the biochemical evidence is suggestive as the prokaryotic proteins lack the negatively charged region postulated to interact with steroid-hormone receptors.

In vertebrate cells, the association of hsp90 with steroid-hormone receptors and tyrosine kinases suggests it may serve as a type of "molecular chaperone," a function that, in the broadest sense, it may share with other heat-shock proteins (61). Unfortunately, this scheme rests almost entirely on in vitro analyses. Genetic analysis in yeast and *E. coli* has provided important general information but few specific biochemical hypotheses. The high degree of conservation in members of the hsp90 gene family and the similarity in their patterns of expression suggests that their roles are similar in all eukaryotes and, certainly in some respects, in bacteria. Given the many different proteins with which hsp90 is believed to associate, genetic analysis is likely to be complicated by pleiotropic effects. Almost certainly, a combination of biochemical and genetic investigations will be required to decipher all of the functions of hsp90. To such an end, a steroid hormone response has recently been produced in yeast cells by transforming them with the gene for the estrogen receptor together with a estrogen-responsive reporter gene (139).

THE HSP70 FAMILY

HSP70 encodes the abundant heat inducible 70 kd hsp. *HSP70* of most, if not all eucaryotes is a member of a multigene family whose genes are expressed under a variety of physiological conditions. The *HSP70*-related genes isolated thus far are related at the level of greater than 50% identity over their entire length. The evolutionary history of this multigene family is not well understood. However, in some cases (such as the *GRP78*-like genes) the similarities are greater among genes from different species than among genes of the same organism, indicating early gene duplication events and maintenance of the multigene family over time. A number of hsp70 and related proteins present in different cellular compartments and associated with a wide variety of cellular processes have been identified. Studies have revealed biochemical similarities among the related proteins from a single organism, as well as among proteins isolated from diverse organisms. All hsp70 and related proteins bind ATP with high affinity (36, 237, 255); many are very abundant in cells and often found in association with other proteins. Genetic analyses, which have been carried out only in *E. coli* and lower eucaryotes, show that *HSP70* and related genes are essential for growth either at high temperatures or at all temperatures, indicating a critical role in normal cellular physiology for the encoded proteins.

Drosophila melanogaster

During studies with *Drosophila melanogaster HSP70* was found to be a member of a multigene family (95, 230). The family includes 5–6 copies of *HSP70* and one copy of the heat inducible *HSP68* gene. In addition, seven other genes that are expressed during normal growth have been identified and denoted *HSC1–7*. *HSC1–6* have been mapped cytologically to 70C, 87D, 10E, 88E, 50E and 5C, respectively (44; K. Palter, E. Craig unpublished results). Using monoclonal antibodies, the presence of two abundant proteins encoded by these genes, called hsc70 and hsc72, have been identified (111, 162). After heat shock of the whole organism, hsp70 and hsp68 become obvious spots when stained with Coomassie blue (162); however, they never attain a level higher than the normally present related proteins (162). hsc70, which is encoded by *HSC4*, is a cytoplasmic protein heavily concentrated around the nucleus. hsc72 is encoded by *HSC3*, and a number of results indicate that it is analogous to the mammalian glucose-regulated protein, grp78 present in the endoplasmic reticulum.

Cellular localization studies of hsps were first carried out in *D. melanogaster*. After heat shock, hsp70 was found to concentrate mainly within the nucleus and secondarily at cell membranes (223). This translocation is not dependent upon the temperature per se, because concentration in the nucleus is also observed after exposure to a hypoxic environment. During recovery

from heat shock hsp70 leaves the nucleus and is found mainly in the cytoplasm. hsc72 is also found in the nucleus after a heat shock (K. B. Palter, G. Gorbsky, G. Borisy and E. A. Craig, unpublished results).

Saccharomyces cerevisiae

The *S. cerevisiae* genome contains at least nine genes related to *HSP70* of higher eucaryotes. Eight of these genes, originally named *YG100-YG107*, have been renamed on the basis of structural and functional similarities: SSA1–4 (stress seventy family *A; YG100, YG102, YG106, YG107*, respectively); *SSB1* and *SSB2* (*YG101* and *YG103*, respectively); *SSC1* (*YG104*); and *SSD1* (*YG105*). Recently, another member of this family, the *KAR2* gene, has been identified (M. Rose, personal communication). The sequence relationships among these genes are complex (see Figure 2), with nucleotide identities ranging from 50–96%. The expression of the family members is modulated differently in response to changes in growth temperature. *SSA3* and *SSA4* are expressed at very low levels during steady-state growth at 23°C, but their expression is greatly enhanced upon upshift to 39°C (240). *SSA2* expression changes little upon shift to a higher or lower temperature. *SSB1* and *SSB2* transcripts are abundant during steady-state growth but rapidly

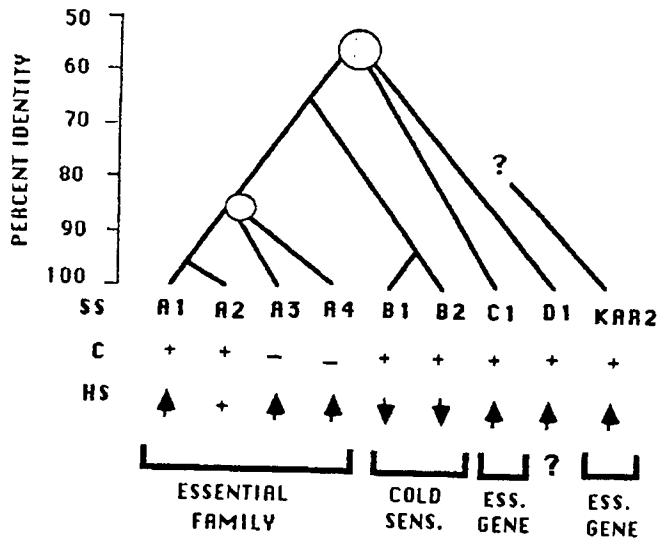


Figure 2 HSP70 multigene family of *S. cerevisiae*. Approximate percentage of nucleotide identities are based on partial or complete sequence data. The expression of the genes during exponential growth at 23°C (C) and after shift to 39°C (HS) and the placement of the genes into functional groups is compiled from Craig et al (47), Ellwood & Craig (62), Werner-Washburne et al (240), Jacobsen & Craig (45, 46), and M. Rose, personal communication.

decrease upon an upshift in temperature (46). *SSA1, SSC1, SSD1*, and *KAR2* transcripts are abundant during steady-state growth and increase 3–10 times upon an upshift in temperature (47, 62; Mark Rose, personal communication). Three members of the family have been precisely mapped. *SSA1* is located on the left arm of chromosome I, 7kb from the centromere (48); *SSC1* is on the right arm of chromosome X, 4–5 kb centromere proximal to *CYC1* (47); *KAR2* is on the left arm of chromosome X (M. Rose, personal communication).

Strains have been constructed containing mutations in members of the HSP70 family. Of the nine genes isolated, *SSD1* is the only member that has not been shown to be functional; no phenotype has been associated with the absence of the gene product. *SSC1* and *KAR2*, both of which are essential genes, are the only members whose absence has been found to result in a phenotypic effect in the absence of mutations in other genes. The *SSB* and *SSA* genes form two additional functional groups. Thus, at least four genetically identifiable functional groups comprise the *HSP70* gene family.

The most complex structural and functional subfamily includes *SSA1-4* (45, 240). Each of the protein products of these genes can substitute at least partially for the absence of the other three. *ssa1 ssa2* mutants are temperature sensitive for growth; they grow at 23°C but are unable to form colonies at 37°C. In addition, they are constitutively thermotolerant, able to survive short periods of time at high temperatures without a prior incubation at moderate temperatures. These cells behave as if they are permanently stressed, synthesizing at high level many heat shock proteins, including hsp90, hsp26 and Ssa4p. The reason for this production of stress proteins is not clear; either the cells are responding because they are "sick" due to the absence of Ssa1p and Ssa2p, or Ssa1p and Ssa2p are part of an important regulatory loop that is broken in their absence. Strains containing mutations in *SSA4, SSA3* or *SSA3* and *SSA4* were found to be indistinguishable from wild-type; they have the same growth and thermotolerance properties. Similarly, *ssa1 ssa2 ssa3* strains behaved like ssa1ssa2 strains. However, *ssa1 ssa2 ssa4* strains are not viable; spores of this genotype do not bud, and vegetatively growing cells containing an *SSA1* gene under the control of a conditional promoter undergo, on average, three cell divisions before growth stops after the termination of *SSA1* transcription. These results indicate that *SSA4*, which is constitutively expressed in the ssa1 ssa2 mutant, is allowing growth at low temperatures. Ssa3p is not synthesized in high amounts in the *ssa1 ssa2* mutant. To test whether *SSA3* could rescue the *ssa1 ssa2 ssa4* cells, *SSA3* protein coding region was put under the control of the constitutive *SSA2* promoter. This construction was able to rescue the growth of *ssa1ssa2ssa4* cells at 30°C, indicating that *SSA3* encodes a protein functionally similar to that encoded by *SSA1, SSA2* and *SSA4*. The reason for the inability of *SSA4* and *SSA3* to allow

growth at 37°C in the absence of *SSA1* and *SSA2* is not clear. Two-dimensional gel analysis of proteins indicate that Ssa4p or Ssa3p is present at levels similar to that of the Ssa1p and Ssa2p in wild-type strains or in strains containing the *SSA2* promoter-*SSA3* structural gene fusion. However, overproduction of Ssa3p or Ssa4p does allow some growth at 37°C. These results suggest that either *SSA3* and *SSA4* proteins are functionally different or predominately in a different cellular location from *SSA1* and *SSA2* proteins.

A strain carrying mutations in either *SSB1* or *SSB2* is indistinguishable from wild-type (46). However, a strain containing insertion mutations in both genes is relatively cold-sensitive for growth. It has an optimal growth temperature of 37°C, growing nearly as well as wild-type at that temperature but 2.4-fold slower at 19°C. A *SSA1* gene under the control of the *SSB1* promoter did not rescue the cold-sensitive phenotype of the *ssb1 ssb2* mutant. Conversely, the *SSB1* coding sequences placed under the control of the *SSA2* promoter could not rescue the temperature sensitivity of the *ssa1 ssa2* mutant. Therefore, the *SSA* and *SSB* genes must encode proteins that are either functionally distinct or present in different cellular locations such that they can not compensate for each other.

Analysis of the *SSA* mutants indicates that these proteins are involved with the posttranslational import of at least some proteins into the endoplasmic reticulum and into mitochondria (53). *ssa1 ssa2 ssa4* cells that contain *SSA1* coding sequences under the control of the *GAL1* promoter cease producing *SSA1* upon shift to glucose-based media. After several hours these cells begin to accumulate precursors to alpha factor (a secreted protein), carboxypeptidase Y (a vacuolar protein), and the beta subunit of F1 ATPase (a mitochondrial protein). Cell-free extracts prepared from *ssa1 ssa2 ssa4* cells after shift to glucose-based media are defective in the import of alpha factor precursor into the microsome vesicles. Furthermore, proteins purified from cells on the basis of their ability to facilitate import of alpha-factor in an in vitro system have been identified as *SSA1* and *SSA2* (38). Thus, both biochemical and genetic evidence indicate that *SSA* proteins are involved in the transport of proteins across membranes. It has been known for some time that an early step in the import of proteins into mitochondria is ATP dependent (59, 170), and evidence suggests that protein must be unfolded in order to pass through the membrane (225). No information exists at this time as to the actual role of SSA proteins in transport. *SSA* proteins may be acting in the cell as an "unfoldase," altering the conformation of proteins in an ATP-dependent manner that allows passage through membranes. Alternatively, they may be involved in some other aspect of the transport machinery and not interact directly with the translocated protein.

SSC1 is an essential gene. *ssc1* spores generated from a heterozygote germinate and undergo approximately three cell divisions before arrest (47).

The predicted amino acid sequence of *SSC1* has an additional 27 N-terminal amino acids when compared to *SSA1* protein. This proposed leader has structural features very similar to those found in proteins imported into mitochondria. The *SSC1* leader is rich in positively charged amino acids, devoid of acidic amino acids, and rich in the hydroxylated amino acids, serine and threonine. Furthermore, *SSC1* protein synthesized in vitro is imported into mitochondria and cleaved to a mature form in an in vitro assay (J. Kramer, M. Werner-Washburne, E. A. Craig, unpublished results). Presumably Ssc1p is a mitochondrial hsp70 protein.

Strains containing *KAR2* mutations were originally isolated because of a defect in nuclear fusion (173). Sequence analysis obtained from a complementing clone revealed a similarity to *HSP70* genes, especially *GRP78*, a mammalian member of the *HSP70* family found in the endoplasmic reticulum. The predicted *KAR2* protein has a leader sequence similar to those found in proteins imported into the endoplasmic reticulum. Furthermore, as expected of a homologue of GRP78, *KAR2* is induced by the glycosylation inhibitor tunicamycin. The *KAR2* gene is essential as shown by gene disruption experiments (M. Rose, personal communication). The reason for the defect in karogamy is not clear. The nuclear envelope may be dependent on proteins and protein complexes that enter and are assembled in the endoplasmic reticulum.

The genetic analysis in yeast leads to the conclusion that there are at least four genetically distinct groups of hsp70-related genes. This distinction could be due either to basic functional differences or to differences in cellular location. It is likely that both explanations are correct. Recent results indicate that at least two of the genes, *KAR2* and *SSC1*, produce products that are localized in the endoplasmic reticulum and mitochondria, respectively. Since both *SSC1* and *KAR2* are essential genes, the encoded proteins likely perform essential functions in these organelles. On the other hand, the genetic results with the *SSB* and *SSA* families indicate that there are functional differences amongst the proteins as well.

Escherichia coli

dnaK encodes a protein that is related to hsp70 of eucaryotes (11); DnaK is 50% identical in amino acid sequence to hsp70 of eucaryotes. There appears to be no other HSP70-related genes in the *E. coli* genome. *dnaK* was originally identified as a host gene necessary for lambda DNA replication (73, 206). Genetic data indicates that DnaK interacts with the P protein of lambda, since mutations in lambda phage that enable it to grow in a *dnaK*$^-$ host map in the *P* gene. Biochemical experiments have confirmed and further defined its role in lambda DNA replication. Six proteins are required for the localized unwinding of duplex DNA at the origin of replication, prior to the binding of

DnaG primase: two lambda proteins—O and P—and four host proteins—DnaB (a helicase), DnaK, DnaJ and Ssb (single stranded binding protein). A complex of O, P, and DnaB form at the origin (117, 118, 246). The subsequent addition of DnaJ, DnaK, and Ssb proteins plus ATP results in an origin-specific unwinding of the DNA duplex. It is hypothesized that DnaJ and DnaK "loosen" the association between DnaB and P, so that DnaB is able to function as a helicase (246), thus permiting DnaG binding and subsequent DNA synthesis (see Figure A).

The *dnaK* gene was so named because *E. coli* DNA synthesis shuts off when mutant cells are shifted to high temperatures. Until recently, however, there was no direct evidence that DnaK was involved in host DNA replication. Sakakibara (184) isolated a new *dnaK* allele, *dnaK111* during a screen designed to isolate mutants defective in the initiation of DNA synthesis. It was known that mutations in *dnaA* (a gene required for DNA initiation) were suppressed in the presence of mutations in *rnh* which encodes RNAase H (93, 124). The suppression is thought to be due to the presence of some latent replication origins, which become active in the absence of RNAase H activity.

dnak111 was isolated as a temperature-sensitive mutant whose defects in

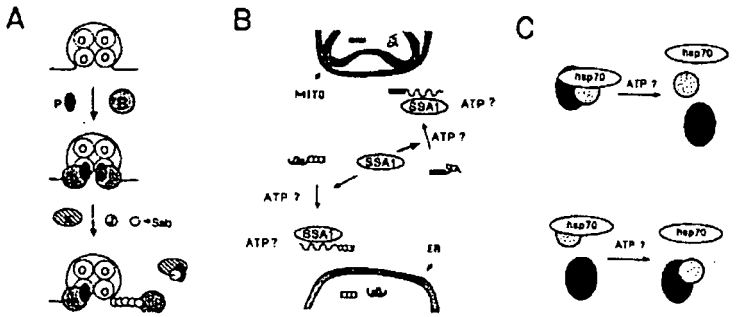


Figure 3 Schematic diagrams of hsp70 function. A. A proposed scheme for the role of DnaK and DnaJ in the initiation of phage lambda DNA replication. In step1, O protein binds to the lambda origin; in step2, P protein and DnaB protein bind; in step3, DnaK and DnaJ bind, disrupting the interaction between P and DnaB and thus allowing DnaB to catalyze the unwinding of the DNA and the subsequent binding of single stranded DNA binding protein Ssb. Adapted from Echols (58). B. Possible role of Ssa proteins in protein translocation across the endoplasmic reticulum (ER) and mitochondrial (MITO) membranes. The interaction with precursor proteins could either be direct as depicted or involve interactions with other components of the translocation machinery. Adapted from Deshaies et al (53). C. A general model of hsp70 and related protein function. hsp70 may be involved in disruption of already established protein-protein interactions between proteins, possibly in an ATP dependent manner (top), or in facilitating the establishment of "proper" interactions (bottom).

DNA synthesis could be relieved upon inactivation of the *rnh* gene. The *dnak111* mutant is unable to initiate a new round of DNA replication at high temperature after termination of the round in progress. DNA synthesis in both the *dnaK111* and *dnaA* mutants becomes temperature independent after reintroduction of a wild-type *rnh* gene. Unlike *dnaA* mutants, however, *dnaK111* mutants in the presence of an inactive *rnh* gene remain temperature sensitive for growth. The inability of *dnak111* to grow at high temperatures even in the absence of *rnh* function suggests that this *dnaK* mutation causes pleiotropic effects. The defect in DNA synthesis can be corrected by inactivating *rnh* but other defects that cause cell death are not. These results imply that DnaK carries out multiple important functions in the cell.

Other genetic analyses have shown that *dnaK* is essential at high temperatures, and perhaps at lower temperatures as well. Cells containing deletions of *dnaK* can not grow at 42°C (161; P. Kang, E. A. Craig, unpublished results). The results at lower temperatures are more equivocal, because it appears that *dnaK* null mutants quickly acquire secondary mutations that allow more vigorous growth (P. Kang, E. A. Craig, unpublished results; A. Bukau, G. Walker, personal communication). Original transductants are extremely filamentous; cells containing secondary mutations are less so. Since cells that do not undergo this change are incapable of continued growth at 30°C or grow very poorly, it is suggested that *dnaK* is necessary for growth at temperatures other than 42°C.

The filamentation caused by *dnaK* null alleles is suppressed by plasmids that carry the *ftsZAQ* genes, which are necessary for normal cell division (A. Bukau, G. Walker, personal communication). The notion that *dnaK* may be involved in cell division, either directly or indirectly, is also supported by the finding that *fam-715* and *rpoH* genes are allelic (214). *fam* has been identified as a gene affecting cell division; *rpoH* encodes the heat-shock specific sigma factor and thus regulates the transcription of *dnaK* and other heat-shock genes.

It is not clear why strains containing *dnaK* null alleles die. Although DnaK is involved in the initiation of DNA replication, its role in the process does not seem to be essential. However, second site mutations that allow growth of *dnaK* null alleles have been isolated (161; P. Kang, E. A. Craig, unpublished results). In addition, overexpression of a wild-type gene allows cell growth at 41°C. This gene, which appears to be previously unidentified, maps to 4 minutes on the *E. coli* chromosome (P. Kang, E. Craig, unpublished results). Analysis of such suppressors may lead to an understanding of the essential functions of *dnaK*.

Dnak has been purified to near homogeniety, and the purified protein has been found to bind tightly to ATP and to have a weak DNA-independent ATPase activity (254, 255). Only 15–20 nmol of ATP are hydrolyzed to

ADP and Pi per milligram of protein per minute at 30°C, representing a turnover number of only one ATP molecule per minute. The DNA-independent nature of the ATPase activity was somewhat unexpected, since a number of proteins involved in DNA replication which also possess ATPase activity show a DNA dependence of the reaction, (e.g. as DnaB, protein m, helicase 1, and helicase 2). The purified protein is capable of self-phosphorylation on a threonine residue. In vivo about 5% of the total cellular DnaK is present in the phosphorylated form after labeling with ^{32}P. The DnaK protein also possesses a 5'-nucleotidase activity which is inhibited by AppppA (20). In vivo analysis of temperature sensitive *dnaK* mutants revealed that five proteins, normally phosphorylated, do not become so at the nonpermissive temperature. Two of the proteins, phosphorylated on threonine residues, have been identified as glutamine tRNA synthetase and threonyl tRNA synthetase (229).

Mammalian Cells

Four members of the human hsp70 protein family have been identified. These proteins have been called by various names in the literature (see Table 1). Here we refer to the human proteins as hsp70, hsp72, p72 and grp78, the nomenclature used by B. Watowich & R. Morimoto (234). hsp70, the major heat-inducible protein (236) is also a cell cycle regulated protein (140). Furthermore, it is under the control of adenovirus E1A protein and is often referred to as the 72k heat-shock protein. hsp72 is a protein which is expressed only after heat shock. p72 is expressed at high levels in growing cells

Table 1

Protein[1] name	Other name(s)	pI	Regulation
hsp70	72K[2]; hsx70[3]; SP71[4]; hsp68[5]	5.8-6.3	Major heat-inducible 70K; basal expression; serum stimulated; cell cycle regulated; E1a inducible
hsp72	hsp70	5.6-5.8	No basal expression; heat inducible
p72[6]	73K[2]; hsc70[5] hsc73[4]	5.5-5.6	High basal expression; slightly heat inducible
grp78	BIP; hsp80	5.2-5.3	High basal expression (especially in secretory cells; expression enhanced by glucose deprivation, calcium ionophores, glycosolation inhibitors, etc.

1. Using the nomenclature of Watowich & Morimoto (234) 2. Welch & Feramisco (236); 3. Pelham (165); 4. White & Currie (241); 5. Lowe & Moran (130); 6. a major heat-inducible protein in primates, but not found in rodents.

and is often referred to as the 73k heat-shock protein. grp78 is a glucose regulated protein located in the endoplasmic reticulum (146). Rodent cell lines synthesize three members of the hsp70 family; there appears to be no direct equivalent of p72 (that is, a protein present in cells at a high basal level and also induced dramatically upon heat shock). A glucose regulated protein, grp75, present in mitochondria appears to be hsp70-related (W. Welch, personal communication). The total number of *HSP70* related genes (proteins) in mammalian cells is not clear. Most of the analyses have been carried out in tissue culture cells. Transcripts from two previously unidentified *HSP70* related genes have been found in mouse spermatogenic cells (2, 248, 249). After careful examination of tissues during development more members may be found. At least 10 *HSP70* related genes have been found in the human genome, but the number that are functional genes is not clear (144). Human *HSP70* related genes have been shown to be located on chromosomes 6, 14, 21, and at least 1 other chromosome (79, 88). As discussed below, biochemical analysis of the mammalian hsp70 related proteins has provided much information.

CLATHRIN UNCOATING ATPase Clathrin uncoating ATPase was identified as a member of the hsp70 family based on the copurification of the uncoating ATPase and p72, two-dimensional gel analysis and immunological cross-reactivity (36, 216). The uncoating ATPase was purified on the basis of its ability to release clathrin triskelions from bovine brain coated vesicles (187). Coated vesicles mediate selective intracellular membrane transport. Clathrin, which is the major structural component of the coated vesicle basket, is found as a three-legged structure called a triskelion. Clathrin triskelions can spontaneously self-assemble into "cages" which resemble the coats of coated vesicles. The uncoating enzyme hydrolyzes ATP in a clathrin-dependent manner, driving cage disassembly (23). During an uncoating reaction, clathrin triskelions are released intact with the uncoating protein in a stoichiometric complex; 3–4 molecules of ATP are hydrolyzed per triskelion released. A two-stage model for the mechanism of clathrin cage disassembly has been presented (189). First, the energy of ATP hydrolysis drives the transient displacement of a portion of a triskelion from the cage. This transient displacement is thought to reveal a previously buried site, which is then bound to the uncoating protein, thus stabilizing the displacement. The triskelion-uncoating complex is released when all points of attachment of the cage are broken (190). In the in vitro assays, the uncoating protein bound to triskelions can be recycled. Since the apparent affinity of uncoating protein for cages is five times higher than that for unassembled triskelions, recycling is a spontaneous process. However, when the reaction is carried out with coated vesicles isolated from cells, the reaction is stoichiometric, that is, a 70K

protein is involved in only one round of uncoating (81). Free clathrin does not inhibit the reaction (81). The uncoating ATPase is not permanently inactivated. If separated from clathrin, it is capable of participating in uncoating again.

At the present time, however, there is no evidence that p72 functions as an uncoating ATPase in vivo. It is unlikely that clathrin uncoating is the only function of p72. In some tissues there is a 30-fold molar excess of p72 over clathrin (76). Mouse stem cells, for example, express p72 at very high levels, but clathrin is barely detectable.

P72 AND HSP70: BINDING TO NUCLEAR MATRIX AND NUCLEOLI Upon heat shock, both p72 and hsp70 migrate to the nucleus and are associated with the "insoluble matrix" in a salt resistant manner (238). The proteins subsequently become associated with nucleoli. Nucleoli are particularly sensitive to hyperthermia; severe structural changes which persist for several hours after a heat shock can be observed microscopically. The recovery of nucleolar morphology occurs more rapidly in cells that are constitutively overexpressing Drosophila hsp70 due to transformation with the Drosophila *HSP70* gene on a plasmid (165). The interaction of p72 and hsp70 prior to heat shock appears to be weak and readily reversible since they are released from nuclei upon lysis of cells with isotonic buffer. After a heat shock, the association becomes strong; however, they are rapidly released in vitro in the presence of ATP (121). Nonhydrolyzable ATP analogues are not effective in affecting release.

INTERACTION OF P72 AND HSP70 WITH CELLULAR TUMOR ANTIGENS Cells transfected with the gene encoding p53, a cellular oncogene, and an activated *ras* oncogene become morphologically transformed. In such cells, p53 is more abundant and stable than in normal cells, and it is associated with p72 (92) and with a lower affinity, with hsp70 as well (171). Mutations in the gene encoding p53 that activate its transforming potential also result in the synthesis of mutant proteins which show preferential association with p72 and have an increased half-life (70). It has been hypothesized that p53 is stabilized because of its interactions with p72. This hypothesis is based on the precedent that the apparent stability of p53 in SV40 transformed cells is caused by its association with the large T antigen. The p72-p53 complex can be dissociated in vitro with ATP, but not with unhydrolyzable analogues. Interestingly, p53 synthesized in *E. coli* is found in association with DnaK (40).

In addition, a mutant medium T antigen encoded by a nontransforming mutant of polyomavirus is reportedly associated with p72, whereas medium T antigens of wild type and at least one transformation-competent mutant polyomavirus are not (232).

GRP78 The grp78 protein, a member of the hsp70 family (146), was originally identified as a protein whose rate of synthesis increased when cells were starved for glucose (119, 193), and it was later shown to be induced under a variety of other conditions including anoxia, paramyxovirus infection and treatment of cells with glycosylation inhibitors or calcium ionophores, but not by heat shock. An inverse correlation between the rate of glycosylation and the steady-state level of the GRP78 transcripts has been observed (35). grp78 has a hydrophobic amino terminal leader sequence which by analysis of gene fusions has been shown to be competent for transport into the endoplasmic reticulum. grp78 is very similar or identical to BiP (146), a protein originally reported to associate with immunoglobulin heavy chains in preB cells that do not make light chains. In normal B cells and plasma cells, a smaller fraction of the intracellular heavy chain is also associated with BiP (87). BiP can be released from BiP-immunoglobulin complexes in vitro by incubation with ATP (146). grp78 (BiP) is present in the lumen of the endoplasmic reticulum of a large number of different types of mammalian cells. It binds transiently to a variety of wild-type secretory and transmembrane proteins and permanently to proteins that are unfolded or misfolded. For example, mature hemagglutinin of influenza virus is a trimeric glycoprotein. Mutants of hemagglutinin that fail to be transported from endoplasmic reticulum to the Golgi apparatus are not efficiently assembled into trimeric structures and remain in a partially unfolded state, associated with grp 78 (75). Mammalian cell lines which have decreased amounts of grp78 expression have been constructed using antisense RNA (54). These lines show increased secretion of mutant proteins.

It is interesting that a 30 amino acid peptide isolated from rat cells with complete identity with the carboxyl terminus of grp78 has been identified as a steroidogenesis-activator polypeptide (164). This peptide, whose activity is increased in the presence of cyclic AMP, is postulated to regulate a commitment step in steroid formation that is under hormonal control, the conversion of cholesterol to pregnenolone. It is thought that this peptide is derived from a larger precursor, presumably grp78 or a closely related protein.

OTHER CHARACTERISTICS OF HSP70 AND RELATED PROTEINS Rat p72 and hsp70 are associated with nonsterified fatty acids, palmitic acid, stearic acid and a small amount of myristic acid (86). The role of fatty acids in hsp70 function is not clear. A number of covalent modifications of hsp70 and related proteins have been reported in the literature. Some of the mammalian proteins are methylated at lysine and arginine residues (233). Dictyostelium hsp70 has been reported to be phosphorylated (129). Attempts to detect phosphorylation of the chicken, mammalian, and Drosophila proteins have not been successful (85; S. Lindquist, K. Palter, E. Craig, unpublished results), however,

phosphorylation of hsc72 of Drosophila has been detected (S. Lindquist, K. Palter, E. Craig, unpublished results)

Discussion of the Hsp70 Function

From the results presented in the previous sections it is obvious that HSP70 and related genes have been implicated in a variety of cellular processes. At first glance these processes—which include DNA replication, transport of proteins across membranes, binding of proteins in the endoplasmic reticulum, and uncoating coated vesicles—appear to have little in common. However, they may all involve the disruption of either intramolecular or intermolecular protein-protein interactions. The model for uncoating of coated vesicles involves the disruption of the coated vesicle basket by disruption of the interactions between the clathrin triskelions in an ATP dependent manner. The hypothesized role of DnaK in lambda replication is the disruption of the interactions of DnaB and P proteins such that DnaB is able to function as a helicase. It has been hypothesized that the role of the SSA yeast proteins in protein import is to disrupt intramolecular interactions of proteins such that they can attain an import-competent conformation. Varshavsky and colleagues (71) suggested that heat-shock proteins might bind to denatured or abnormal proteins after a heat shock to prevent their aggregation and thus to prevent cellular damage. Pelham (166) extended this hypothesis to include the assembly and disassembly of proteins and protein-containing complexes both during normal growth and after a heat shock. As noted above, hsp70 and related proteins all have a high infinity of ATP. Most of the "reactions" described involving 70K proteins, including the release from nucleoli require, or are at least thought to require, the hydrolysis ATP. Probably the disruption of the protein-protein interactions require the energy generated from ATP hydrolysis. The findings that hsp70-related proteins reside in the endoplasmic reticulum and mitochondrium and that hsp70 is translocated into the nucleus upon heat shock suggest that 70K proteins perform important functions in all cellular compartments.

GRO E-HSP58

The *groE* genes of *E. coli* were originally identified as genes necessary for productive growth of bacteriophage lambda and T4 (74). The two genes *groEL* and *groES* comprise an operon under heat shock control (212) located at 93.5 minutes on the *E. coli* chromosome (84). They encode abundant proteins; *groEL* a 65 kd Mr protein, *groES* a 15 kd Mr protein. In its native form *groEL* protein is a decatetramer, with its subunits arranged in a double ring with seven-fold symmetry (90).

Although *groEL* and *groES* are essential for cell growth at all temperatures

(C. Georgopoulos, personal communication), their role in cell growth is unclear. Temperature-sensitive *groEL* and *groES* mutants show an inhibition of both cellular DNA and RNA synthesis at the nonpermissive temperature (211, 228). Suppressors of one *groES* allele map to the *rpoA* gene that encodes a subunit of RNA polymerase (227). Furthermore, overproduction of GroEL and GroES suppress some *dnaA* alleles. *dnaA* is a gene essential for the initiation of *E. coli* DNA synthesis (66, 97). Since RNA polymerase is necessary for priming of DNA synthesis, these genetic results implicate the *groE* genes in some aspect of DNA replication.

GroEL and GroES are required for head assembly of lambda and T4 phages and for tail assembly of T5 phage. The effect of *groES*$^-$ mutations on lambda head assembly can be suppressed by mutations in *groEL* suggesting an interaction of the encoded proteins in vivo (210). Results of biochemical experiments suggest an association as well. The two proteins cosediment in a glycerol gradient in the presence of ATP and Mg^{2+}, and GroES binds to a GroEL-affinity column (34). Analysis of *groEL* mutants indicates that GroEL acts in lambda prophage head assembly at a step involving the bacteriophage-coded minor head protein B, in which B protein is oligomerized into a dodecameric structure (107). The B protein decamer is located on the head at the point of tail attachment.

Recently a heat inducible protein, (called hsp58), which is related to GroEL, has been identified in Tetrahymena (136, 137). hsp58 is constitutively expressed, and its level increases two-to-three-fold after heat shock. The majority of hsp58 is mitochondrial associated and in a nondenaturing gradient sediments as a 20–25S complex. Proteins that cross-react with an antibody directed against the Tetrahymena protein have been identified in a wide variety of species, including yeast, frogs, maize and human cells. Furthermore, a very abundant chloroplast protein called the Rubisco large subunit binding protein is related to GroEL (89). A large subunit of rubisco (ribulose biphosphate carboxylase-oxygenase) contains the catalytic site but is active only in oligometric combination with the small subunit. The GroEL-like protein is implicated in the assembly of the Rubisco multimetric complex. Similarities between this role in chloroplasts and the role in phage morphogenesis can be seen; both involve the assembly of multimetric complexes. The *groE* proteins may possibly play a similar role in DNA replication as well.

THE SMALL HSPS

The small hsps are a very diverse group. Different organisms have different numbers of small hsps, ranging from one, in *S. cerevisiae* (168), to upwards of 30 in higher plants (131). Great variance is also observed in molecular

weights, which range from 16 kd in the nematode *C. elegans* (182), to 40 kd in the protozoan *S. mansoni* (150). Even within species, considerable heterogeneity is observed (131). Nevertheless, the small hsps of different organisms are clearly related. They have similar hydropathy profiles (although such analyses must be treated with caution) and small regions of amino acid identity (See Figure 4). Their most invariant feature is the sequence aa1-aa2-glycine-aa3-leucine-aa4-aa5-aa6-aa7-proline-aa8, which is found near the carboxy terminus at a similar position in the hydropathy profiles of each protein. The amino acids designated aa1-8 are not invariant but are represented by only a small number of amino acids. For example, aa4 is threonine in most species; aa5 is always isoleucine, valine, or leucine. It is striking that the small hsps show much greater homology within organisms than between organisms. For example, members of a subgroup of the soybean small hsp family have 90% amino-acid identity with each other but only 20% amino-acid identity with the proteins of *D. melanogaster*, *C. elegans*, and *X. laevis* (148). Either the genes are subject to relatively frequent conversion events, or they have been frequently deleted and expanded during evolution. None of the well-characterized hsps of *E. coli* show significant homology to the small hsps of higher cells, and until recently, it was believed the protein family was eukaryotic in origin. However, the gene for an 18 kd antigen that shows homology with the eukaryotic small hsps has now been isolated from the mycobacterium *M. leprae* (247) (see Perspectives, above). It now appears that the small hsp family has existed for well over a billion years.

Though divergent in sequence, the small hsps are conserved in their structural properties. First, they share the property of forming highly polymeric structures, often called heat-shock granules, with sedimentation coefficients of 15–20S (9, 157, 191). As first reported for the fruit fly *D. melanogaster*, the small hsps have sequence homology with the vertebrate alpha-crystallins, which form similar structures (19, 96). Unfortunately, this characteristic has led to some confusion, as other particles of similar size, also composed of small molecular-weight proteins, are found in most eukaryotic cells (7, 188). These other particles, called prosomes, or proteasomes because of their involvement in proteolysis, have recently been distinguished from the heat-shock granules by three different research groups (8, 64, 157). The small hsps appear to share other physical properties as well. Both the mammalian proteins (9, 103) and the *Drosophila* (178) proteins are phosphorylated under a variety of different conditions. It is not yet known whether this modifiation is universal or serves any purpose. Finally, the small hsp particles isolated from *Drosophila*, sea urchin, and tomato cells (106, 157, 176) have been found to contain RNAs. The protein may serve to preserve translationally inactive messenger RNAs (157). However, this finding has been disputed (41), and mutations in the unique small hsp of yeast cells (see below) have no

affect on the stability of stored mRNAS in yeast spores (S. Kurtz, S. Lindquist, unpublished). Again, that this property is shared by such evolutionarily diverse organisms suggests it is significant, but its function is as yet unclear.

Finally, the small hsps share the property of being induced at specific stages in development at normal temperatures. These developmental induction patterns can be quite complex. In *D. melanogaster*, a total of seven small hsps

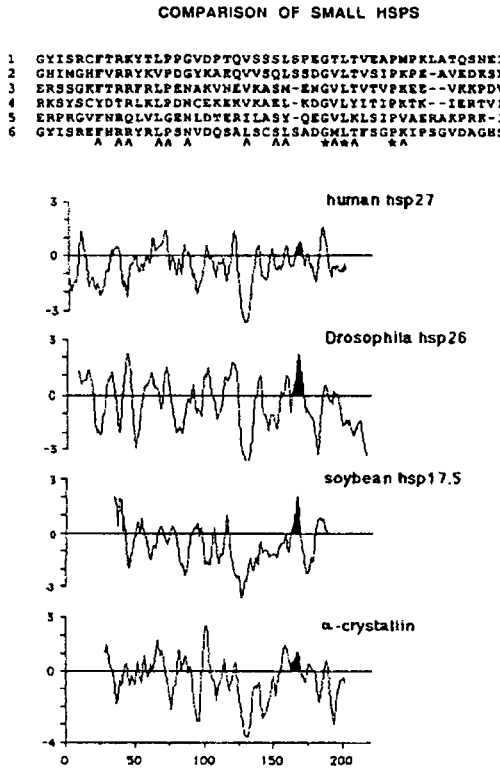


Figure 4 Comparison of the small hsps.
Top: the amino acid sequences of the most conserved domain, located at or near the carboxxy terminus of several small hsps. 1. Human hsp27 (aa # 132–186) Hickey et al (91). 2. Drosophila hsp26 (aa # 127–180) Ingolia & Craig (96). 3. Soybean hsp17.5E (aa # 103–154) Czarnecka et al (50). 4. Pea chloroplast hsp21(aa # 183–232) Vierling et al (226). 5. Mycobacterium 18kd antigen (aa # 81–133) Nerland (151). 6. Bovine α-crystallin A chain (aa # 108–161) van der Ouderaa (217). (*) Invariant amino acids; (˙) identical amino acids in at least 4 of the 6 proteins. Bottom: hydropathy profiles of members of the small hsp protein family (as analyzed by the method of Kyte & Doolittle, (112), with a window of 6). The profiles are aligned by the most conserved domain, which is indicated in black.

have been identified, six of which clearly belong to the small hsp family (10, 96, 201). All of the genes are transcribed during the late larval and early pupal stages (10, 37, 198). *GENE 3* is expressed in the middle of embryonic development (163). *HSP 23* and *GENE1* are transcribed in young adults just after eclosion (10). *HSP27* and *HSP26* are expressed in nurse cells (together with *HSP83*) and are passed into developing oocytes (252). Detailed, tissue-specific analysis of one of these genes, *HSP26*, revealed further complexities, with expression concentrated in spermatocytes, nurse cells, epithelial tissues, imaginal discs, proventriculus, and neurocytes (78). At least some of this stage- and tissue-specific regulation is due to the hormone ecdysone (15, 208), which exerts its effect on transcription elements separate from the heat-shock elements on the small hsp genes. The elements that regulate transcription in spermatocytes can also be separated from those regulating transcription in nurse cells. These findings suggest that a highly complex regulatory system is operating on these genes.

Although less extensively characterized in other organisms, it seems likely that developmental induction of the small hsps will be universal. In *Saccharomyces* hsp26 is induced during stationary phase growth and during sporulation (111). In lilies several small hsps are induced during meiosis (R. Bouchard, personal communication). The presence of such developmental induction in these diverse organisms implies that small hsps play a role in development as well as in response to stress.

Unfortunately, the molecular role of these proteins is at present a complete mystery. Several studies have suggested that the small hsps are responsible for acquired thermal tolerance. For example, in *D. melanogaster*, ecdysone induces both thermotolerance and the small hsps, but not hsp70 (15). Also, thermosensitive tomato cells constitutively synthesize hsp68 and hsp70, but they become thermotolerant when given a mild heat treatment that induces the small hsps (156). A possible problem with these studies is that small changes in other proteins might have important biological consequences and yet escape detection. Two lines of genetic analysis also support this view. First, a mutant strain of *Dictyostelium* that fails to synthesize the small hsps also fails to acquire thermotolerance (129). However, since the primary lesion in this mutation has not been identified, it may be the failure to acquire thermotolerance that is, in turn, responsible for the failure to synthesize the small hsps. More recently, variants of a CHO cell line, selected for heat-resistance, were found to have elevated levels of hsp28, and only hsp28 (115).

On the other hand, in *Drosophila*, deletion (197) and insertion mutations in individual small hsp genes (60) as well as pseudo-mutations induced by antisense-RNA (135), have no noticeable effects on viability. Of course, in *Drosophila* cells other members of the small hsp family might cover the function of the mutated protein. The genetic analysis in *S. cerevisiae* is more

compelling. In this organism there is only one major small hsp. Deletion mutations have no detectable effect on (*a*) respiratory or fermentative growth at 25°, 35°, 37° or 39°C, the latter being the highest temperature at which these cells will grow, (*b*) basal or acquired thermotolerance in respiring or fermenting cells, (*c*) spore development with or without a heat shock, (*d*) spore germination with or without a heat shock, (*e*) spore viability during long-term storage, (*f*) dessication tolerance, (*g*) ethanol tolerance, (168), or (*h*) growth under anaerobic conditions at high or low temperatures (M. Schlessinger, N. Collier, personal communication). The assays used in these experiments were very sensitive. Large differences in thermotolerance and growth rates in different media and at different temperatures were found in comparing various common laboratory strains of yeast. But, in comparing isogenic wild-type and *hsp26*$^-$ cells, in either high thermotolerance or low thermotolerance genetic backgrounds, no differences were observed. Taking an alternative approach, *hsp26* coding sequences were placed under the control of the *gal1* promoter, to provide induction of the protein prior to heat shock. Cells carrying the wild type *HSP26* gene, the *hsp26* deletion mutation, or *pGAL1-HSP26* showed no significant differences in basal or acquired thermotolerance when grown on galactose (A. Taulien, A. Petko, S. L. Lindquist, unpublished).

Three explanations for the yeast results seem plausible. First, it may be that the function of hsp26 is covered by another protein in *Saccharomyces*. If so, that protein bears little homology to hsp26. No cross-reacting genes have been detected by low stringency DNA hybridization, and no cross-reacting proteins have been observed with a polyclonal anti-hsp26 serum (168; see below). Given the divergence observed in this protein family, other members of the family might well escape detection. However, if other members of the family are present, they must not be strongly heat-inducible, or they would have been detected on gels. A second possibility is that the function of the protein may be much more subtle than previously expected. Genes that give very small advantages for survival under very particular circumstances can still be selected for in evolution. Finally, however unpalatable the notion, it is possible that the small hsp genes represent an ancient form of selfish DNA.

Investigations of the intracellular location of the small hsps have not helped to clarify their functions. Several cell-fractionation studies have localized them to chromatin, nucleoli, and the nuclear skeleton. More recently, immunofluorescence and immunoelectron microscopy has demonstrated that they are in large, cytoplasmic, granular aggregates, often in close proximity to the nucleus (41, 152). During heat shock they partially localize to the nucleus, although this has been disputed (reviewed in 180). In soybeans, at least one member of the small hsp family is transported to chloroplasts and carries a signal sequence that is clipped during transport (226). In the yeast *S.*

cerevisiae, immunofluorescent staining has revealed that the intracellular location of hsp26 has a complex but as yet unilluminating dependence on the physiological state of the cell (180).

At present, the small hsps are the most baffling of all the hsps. They are abundant, ubiquitously distributed proteins induced both by heat shock and by normal developmental cues. They are a remarkably diverse group, conserved more in their structural properties than in their amino-acid sequences. In fact, were it not for their common induction patterns, the homologies in their amino acid sequences would not be sufficient to suggest a common function. Some evidence suggests they are important in thermotolerance; genetic analysis in yeast suggests they play no role in this phenomenon.

UBIQUITIN

Ubiquitin, a highly conserved 76-residue protein, which as its name implies is found in all eucaryotic cells, is induced by heat (21, 22, 72). In all five eucaryotic species examined, it is synthesized as a polyprotein termed polyubiquitin, generally consisting of tandem repeats of the protein coding sequences with no spacer regions. Ubiquitin is found in cells either free or linked via its terminal glycine residue to a variety of cellular proteins. The conjugation process, which is ATP dependent, is apparently an essential precondition for selective degradation of intracellular proteins.

The first mutation identified in a gene involved in the ubiquitination process was in the strain ts85, a temperature-sensitive derivative of a cell line established from a spontaneous mouse mammary carcinoma. This cell cycle mutant has a temperature sensitive defect in the ubiquitin activating enzyme E1 (71). At the nonpermissive temperature (which is below the temperature necessary to induce hsps in mammalian cells) synthesis of ubiquitin and some members of the hsp70 family is increased. This result led to the suggestion that inactivation or overloading of the ubiquitin system leads to the induction of heat-shock proteins (71). It has also been hypothesized that the common induction signal in the heat-shock response is the denaturation of proteins (3). Inducers such as amino acid analogs, heat, and ethanol might cause a fraction of the protein of the cell to denature. The heat shock–specific transcription factor is present predominantly in an inactive form in noninduced cells. The hypothesis is based on the assumption that at any time some of the factor is converted to an active form and that this active form is labile, rapidly inactivated proteolytically, and thus does not accumulate in noninduced cells. When cells are stressed, the protein that is denatured could compete effectively with the active form of the transcription factor for the limited proteolytic machinery, thus resulting in the factor's stabilization. The induction of hsps in line ts85 is consistent with this model. Also, it is known that proteins

containing amino acid analogs are degraded by the ubiquitin system, and that amino acid analogs induce the heat-shock response. However, if this model is correct it is difficult to understand the induction of hsps in situations in which the temperature is raised to only 37°C (a condition which will induce the response in yeast), or to even lower temperatures, as is the case with arctic fishes. It had been proposed that ubiquitin is directly involved in the modification of the heat shock transcription factor and therefore the induction of the response (145). However, no evidence supports this idea. Since it is now known that phosphorylation of the transcription factor occurs upon heat shock, it is believed that this modification is responsible for activation (199).

Ubiquitin genes have been isolated from several organisms; the yeast genes have been studied most extensively. The four yeast genes all encode hybrid proteins in which ubiquitin is fused at its carboxyl end either to itself, as in polyubiquitin encoded by *UBI4* (160), or to unrelated amino acid sequences as in the case in *UBI1*, *UBI2* and *UBI3* (159). Of these four genes only *UBI4* is heat inducible. A precise deletion of *UBI4* has been constructed and substituted into the genome (72). The *ubi4* deletion grows at rates comparable to wild-type strains at least between 23 and 36°C. The steady-state levels of free ubiquitin are very similar to those in wild-type strains, indicating that free ubiquitin is being generated from the products of the other genes. A number of phenotypes of the *ubi4* strains have been observed. *ubi4* strains are hypersensitive to exposure to 38.5°C, a borderline growth temperature for yeast. While about 60% of wild-type cells maintain colony-forming ability for 16 hours at 38.5°, only 1–5% of mutant cells survive. The *ubi4* strains are also more sensitive than wild-type strains to amino acid analogs and starvation for nitrogen and carbon. *UBI4* is also required for maintenance of spore viability. Since ubiquitin is thought to be necessary for the degradation of abnormal proteins, there is probably an increased demand for ubiquitin during times of stress, such as after a heat shock or in the presence of amino acid analogs. Although in a *ubi4* mutant there is sufficient free ubiquitin generated from the products of the *UBI1*, *UBI2* and *UBI3* genes under normal growth conditions, during or after stress there is not; thus, a requirement for *UBI4* under those conditions.

OTHER PROTEINS

Eukaryotes

In various cell types, many other proteins have been found to be heat-inducible. These proteins may be produced in lesser quantities than the classic hsps, may be less strongly heat-inducible, or may be of more limited phylogenetic distribution. They have been little studied and the elucidation of their functions is just beginning.

High-resolution 2-D gel electrophoresis of various species has identified a large number of heat-inducible proteins, which are among the less abundant cellular proteins. One of these, whose rate of synthesis may be induced up to 15-fold, was identified as the α-subunit of eukaryotic initiation factor 2 (eIF-2α). The phosphorylation of eIF-2α was previously proposed as the mechanism for translational repression of normal proteins in heat-shocked mammalian cells (57). Its induction by heat shock might then play a role in restoring normal levels of translation after heat shock. However, the role of eIF-2α in regulating heat-shock translation is still controversial (132), and no data on the generality of its induction by heat are yet available.

The gene for a 35-kd heat-inducible protein in *S. cerevisiae* was identified as one of three closely related genes in the yeast genome (169), previously identified as the glyceraldehyde-3-phosphate dehydrogenase gene family (G3PDH). The isoform induced by heat is also one of the most abundant cellular proteins at normal temperatures. A strain carrying a disruption mutation grows more slowly than the wild-type in rich media at 25 and 37°C, but it grows at the same rate as the wild-type in rich media at 39°C and in minimal media at 25°C. Apparently the heat-inducible G3PDH isoform facilitates rapid growth over a broad range of temperatures. However, when growth rates are reduced, either by high temperatures or by a lack of nutrients, it is dispensable. In media in which the mutant and wild type grow at the same rate, the mutant is more resistant than the wild-type to killing at 50°.

A heat-inducible protein of ~35-kd has been reported in many cell types and may be a common feature of the response in eukaryotes. In most cases the identity of the protein is unknown but in *Xenopus* embryos hsp35 has also been identified as G3PDH (153). It seems likely that its role in the response may relate to the fact that ATP levels are reduced by heat shock (69) while respiration is adversely affected. G3PDH may be induced to help restore ATP concentrations to normal by increasing the rate of glycolysis. That two other glycolytic enzymes, enolase (94) and phospho-glycerate kinase (172), are induced by heat shock lends support to the hypothesis. The induction of enolase (hsp48) requires further note. It was originally reported that yeast cells carrying mutations in the enolase gene are thermosensitive. However, further study of isogenic strains revealed no difference in thermotolerance between *eno1* and *ENO1* strains (H. Iida, personal communication).

Other proteins may be abundant and strongly heat-inducible but idiosyncratic in their appearance, restricted to certain organisms, to certain cell types within organisms, or to certain developmental stages. An example of this type is a collagen-binding protein of chicken embryos (183). In the liver this 47-kd glycoprotein is present in fibrocytes, Kupffer cells and smooth muscle, but it is absent from hepatocytes, bile duct epithelia and sinusoidal endothelium. Other example are (*a*) a 180-kd protein in endothelial cells identified as thrombospondin (102), (*b*) several cuticle-like proteins in epidermal cells of *Manduca*), (*c*) 28 proteins in the male accessory glands of *Sarcophaga*, (100) and (*d*) six heat-inducible proteins produced by fruiting cells but not vegetative cells of *Myxococcus xanthus* (259). This list, though incomplete, is thought provoking. Specialized cells may induce special proteins, due to the particular pathological effects that heat produces in them.

Prokaryotes

Seventeen heat-shock proteins have been identified in *E. coli*. Ten of these are the products of known genes (see F. Neidhardt, R. VanBogelen, 149, for a review). Because of their structural relationship to eucaryotic heat shock genes, three genes, *dnaK*, *groEL* and *htpG*, have been discussed above. Two others, *grpE* and *lon*, are discussed below because of their relationship to *dnaK* and the involvement of proteolysis with the heat-shock response, respectively.

grpE encodes the 24 kd hsp, B25.3 (5). Like DnaK and DnaJ proteins, GrpE is necessary for lambda phage growth at all temperatures, although its role is not clear. Analysis of temperature sensitive mutants indicate that GrpE protein is also essential for cell growth at least at high temperatures (43.5°C and above) (5). The *dnak* and *grpE* proteins interact, as shown both by sedimentation in glycerol gradients and by binding of *GrpE* to a DnaK affinity column. The proteins dissociate in the presence of ATP (253). Extragenic suppressors of the *grpE* 280 mutation have been mapped to *dnaK*, supporting the existence of functional interactions between GrpE and DnaK in vivo (C. Georgopoulos, personal communication).

The ATP-dependent Lon protease is a heat-inducible protein of 94 kd. Mutations in the *lon* gene cause a 2–4-fold reduction in the rate of degradation of incomplete peptides or proteins containing amino acid analogs (39). Also, *lon* mutants have a decreased rate of breakdown of short-lived regulatory proteins, causing a variety of phenotypes (142), including the accumulation of large quantities of mucopolysaccharides, an abnormal SOS response and a decreased ability to lysogenize phage lambda. Overproduction of Lon results in increased rates of degradation of abnormal proteins and normal cellular proteins. The growth of cells containing greater than normal amounts of Lon is also impaired (80). Therefore *Lon* is thought to play a major role in the degradation of abnormal proteins and in regulating the turnover of normal proteins. The fact that both eucaryotic and procaryotic cells have at least one heat-shock protein (the ubiquitin and Lon proteins, respectively) that is involved in proteolysis suggests that one role of the response is the destruction of abnormal proteins accumulated during the stress.

Thermotolerance

The heat-shock proteins are postulated to protect organisms from the toxic effects of heat and other forms of stress. The many mechanisms employed to

ensure that hsps are produced as rapidly as possible after temperature elevation, and the relationship between the temperatures that will induce the proteins in various organisms and the temperature fluctuations of their environment—both support this view. More compellingly, in a remarkable range of cells and organisms, incubation at temperatures that induce the hsps produces tolerance to much more extreme temperatures (122, 148). Examples include vertebrate tissue-culture cells, whole mice, *Drosophila* embryos, larvae, pupae and adults, slime molds, sea urchin embryos and plutei, soybean seedlings, yeast, and bacterial cells. The protective effects of pretreatment are not only manifest in lethality. Sublethal heat treatments induce developmental anomalies in many organisms, including vertebrates, insects, and plants, and preheat treatments reduce or eliminate these defects (141). Of course, many protective changes in physiology might be made during the pretreatment. Is the induction of hsps a critical factor? Here we review experiments that examine the hypothesis. The results of most support it; some do not.

The Correlation Between Hsp Synthesis and Thermotolerance

In experiments measuring the rate of thermotolerance development, it closely parallels the rate of hsp acumulation. Moreover, the decay of thermotolerance when cells are returned to normal temperature, parallels the degradation of hsps. Tolerance can also be induced by other types of conditioning treatments. These have in common the property of inducing hsps. Exposure to ethanol, hypoxia, and heavy metal ions are commonly employed. Such treatments do not induce heat-shock proteins in all organisms, but when they do, they also induce thermotolerance. Significantly, the converse is also true. That is, heat treatments induce tolerance to ethanol, anoxia, and several other forms of stress, underscoring the broadly protective nature of the response. A notable exception is the induction of hsps by amino-acid analogs, which is not accompanied by the induction of thermotolerance. Since proteins that have incorporated amino acid analogs are not likely to be functional, this tends to support, rather than negate, the hypothesis. (Reviewed in 32, 79, 122, 127, 141, 148, 205)

Early in the embryonic development of many organisms, including fruit flies, sea urchins, frogs, and mammals, hsps are not inducible and the organism is hypersensitive to thermal killing. At the time when hsps become inducible, the organism becomes more thermotolerant (reviewed in 25). Sperm development is extremely sensitive to high temperatures. In *Drosophila*, primary spermatocytes do not respond to high temperatures by inducing hsp70 (257). In rodent brains, different cell types exhibit marked differences in hsp70 induction and these correlate with the ability of individual cells to survive ischemia and heat shock (221). In other experiments the induction of hsps has been blocked by inhibitors such as cycloheximide. Usually, the acquisition of thermotolerance is also blocked (127, 134).

There are, however, several counter-examples. In the mouse, spermatocytes do respond to heat by producing hsp70 (2). And a block in protein synthesis in yeast cells does not prevent the induction of thermotolerance by mild heat pretreatments (235). Since yeast cells produce hsps at moderate levels even at normal temperatures, the pretreatment may provide an opportunity to activate pre-existing proteins. Similar results have been obtained in some mammalian cell cultures, and the same argument may apply (120).

Mutations That Alter Hsp Synthesis

In *E. coli*, a mutation originally characterized as temperature sensitive for growth is now known to be an *amber* mutation in the *rpoH* gene, which produces the sigma factor required for expression of hsps (149, 245). When various suppressors are introduced into these cells, their ability to grow at high temperatures correlates with the efficiency of the suppressor. Furthermore, when wild-type strains are transferred from 30°C to 42°C, they acquire tolerance to 55°C, but when mutant strains carrying a ts *amber* suppressor are exposed to 42°C, they do not. Temperature resistant variants isolated from such strains are partially or fully restored for hsp synthesis (213).

In another series of experiments (219), the coding sequences of the *rpoH* gene were placed under the control of other promoters, allowing artificial induction of hsps at normal temperatures with IPTG. IPTG-induced cells showed no increase in thermotolerance, compared to uninduced cells, when they were exposed directly to high temperatures. As with wild-type, they required a conditioning treatment at 42°C, and this treatment was ineffective in the presence of chloramphenicol. A cautionary note is that hsps were not found in the same relative concentrations after IPTG induction as after heat shock, and a few hsps were not induced at all.

It has not been possible to expose eukaryotes to the same tests. The gene encoding the heat-shock transcription factor is essential in yeast (G. Weiderech, H. Pelham, personal communication), and conditional mutations have not yet been produced. However, other yeast mutations do support the argument that hsps are involved in thermotolerance. As discussed above, strains carrying mutations in the *SSA1* and *SSA2* genes, members of the *HSP70* gene family, constitutively overexpress other hsps. When these cells are directly exposed to high temperatures, they are nearly as thermotolerant as the wild-type strain is after a conditioning pre-heat treatment (45). Strains carrying mutations in *HSP83* and *HSC83* show reduced thermotolerance when grown in a medium that supports respiration, but not when grown in media that support fermentation (K. Borkovich, F. Farrell, D. Finkelstein, S. Lindquist, manuscript in preparation). Ubiquitin mutants are hypersensitive to

chronic heat stress, that is, incubation at temperatures that are just slightly above their maximum growth temperature (72). A yeast mutant, *hsr1 (cyr1)*, selected for thermoresistance synthesizes two 48 kd hsps (enolase isoforms) and two other proteins of 73 and 56 kd (94). Finally, yeast cells carrying mutations in various genes that regulate cAMP metabolism fail to respond to nutrient deprivation and are much more sensitive to heat than are wild-type cells (30). This may be because nutrient deprivation normally induces hsp accumulation in wild-type cells.

These experiments provide other important information. First, it is clear that the factors which permit cells to grow at the upper end of their natural temperature range differ from those which permit them to survive short exposure to extreme temperatures. The *ssa1 ssa2* mutant, which has increased tolerance at 50°C, is temperature sensitive for growth (45). The *hsp83* and *hsc83* mutants are also ts for growth, yet have normal basal and acquired thermotolerance when growing by fermentation (K. Borkavich, F. Farrell, D. Finkelstein, S. Lindquist, manuscript in preparation). The latter mutations also suggest that different proteins may be required under different growth conditions, since they have reduced thermotolerance when growing by respiration. The ubiquitin mutations separate the mechanisms required for surviving chronic exposure to superoptimal temperatures (under these conditions they have reduced levels of survival), from those required for surviving short exposure to extreme temperatures (they have normal survival) (72; D. Finley, personal communication). The *hsp35* mutations separate effects on growth rate from thermotolerance. In media in which they grow as well as do wild type, they have higher than normal levels of thermotolerance. In media in which they grow more slowly than wild type, they have normal levels of thermotolerance (169).

Experiments in vertebrate cells complicate the picture further. The results of many experiments implicate hsp70 in thermotolerance, since thermotolerance shows the best correlation with hsp70 concentrations (122). In cells subjected to repeated lethal heat treatments, in order to select thermoresistant variants, the only protein that is constitutively overproduced is hsp70 (122). However, another thermoresistant cell line constitutively overproduces only hsp90 (244), another overproduces only hsp28 (J. Landry, personal communication), and yet another overexpresses hsp89 and hsp68 and a novel hsp70 variant (4). In a complementary series of experiments, cells that have reduced abilities to survive high temperatures are defective in induction of hsp70. On the other hand, some cell lines with widely different levels of thermotolerance exhibit no qualitative or quantitative differences in hsp synthesis (67). Also, mouse cells transformed with SV40 show increased sensitivity to heat, yet both the constitutive and inducible levels of hsps are higher than in the parental line (258).

Overall, a great many experiments support the hypothesis that hsps are important components in the induction of thermotolerance. It seems almost certain that they are. However, in several cases, the synthesis of hsps is not sufficient to provide thermotolerance, and in others, it does not appear to be necessary. There are too many contrary reports, in a variety of different systems, for them to be dismissed. Accepting the premise that hsps do play a role in thermotolerance, two other conclusions seem appropriate: First, cells must have a variety of mechanisms for coping with the toxic effects of high temperatures that are separate and apart from the synthesis of hsps. If the state of metabolism, differentiation, or experimental intervention should prevent these from being activated, the synthesis of hsps might be irrelevant. Second, cells may be killed by different lesions when they are in different states of metabolism or in different stages of differentiation, as when they are exposed to extreme temperatures for short periods or moderately high temperatures for long periods. Different hsps may protect cells from different lethal lesions. Although many of the toxic effects of heat have been defined—reductions in protein synthesis, transcription, and RNA processing, rearrangements of the cytoskeleton, changes in membrane permeability, disruptions in oxidative respiration, photosynthesis, etc.—it is still not clear which of these are critical lethal lesions (reviewed in 181). Defining the specific lethal lesions that are induced by high temperatures would be of great help in defining the mechanisms that are employed for protection. The very nature of the heat-shock response suggests it is homeostatic. The extraordinary degree of conservation observed in most of the proteins indicates the underlying basis for the protective strategies is universal.

REGULATION OF THE RESPONSE

It is difficult to close this review without briefly considering regulation, both because regulation of the proteins is closely associated with regulation of thermotolerance and because studies of heat shock regulation have provided many important insights on the control of gene expression. In *E coli*, the heat-shock response is transcriptionally regulated by the cellular concentration of σ^{32}, a sigma factor that binds to core RNA polymerase and redirects it to heat-shock promoters (82, 114, 245). Simple as it would appear, the regulation of σ^{32} concentrations is itself complex. Transcriptional, translational, and posttranslational mechanisms are all employed. The σ^{32} gene, *rpoH*, is transcribed constitutively by the standard, σ^{70}-containing polymerase. An extremely rapid response to heat is achieved by an immediate increase in the translational efficiency of the σ^{32} message, by an increase in the concentration of the σ^{32} message, and by the stabilization of the normally very unstable protein product (203, 209).

Transcriptional regulation also plays an important role in regulation in eukaryotes. Here the essential transcription factor, HSF (for Heat-Shock Factor), preexists in sufficient concentration but is inactive in form. The factor is rapidly activated in response to temperature elevation by posttranslational modification, producing an extremely rapid increase in heat-shock transcription (199, 243). In higher eukaryotes, additional mechanisms are employed to circumvent other barriers to a rapid response. For example, in organisms with larger genomes and a hierarchy of chromatin structure, the heat-shock genes are preset in an open chromatin configuration at normal temperatures, with hypersensitive sites at their 5' ends (101, 242). In at least some cases, polymerase is already engaged on these genes but is blocked by a negative regulatory mechanism that is immediately released when the temperature is raised (77). Translational mechanisms also play a vital role. While heat-shock messages are translated with high efficiency, preexisting messenger RNAs are translationally repressed, reducing the competition for translation (126, 202). At the same time, hsp70 messenger RNAs, which are extremely unstable at normal temperatures, are stabilized by heat shock (167, 207).

Thus, in both prokarotes and eukaryotes many regulatory mechanisms, acting transcriptionally and post-transcriptionally, are employed to induce the proteins, their unifying theme being to ensure that the proteins are induced as rapidly as possible.

CONCLUDING REMARKS

A theme that runs throughout this review is the involvement of hsps in protein-protein interactions. Small hsps form large aggregates; hsp90 interacts with steroid receptors and with the virus encoded transforming protein, *src;* hsp70 and related proteins with clathrin baskets, DNA replication complexes, ER proteins, and the cellular tumor antigen p53. Much of this data is at least consistent with the notion that some hsps are involved in protein folding and assembly (or disassembly) of protein complexes. The heat inducible proteins may be involved in reassembling structures damaged by heat shock or other stresses.

This review focuses on the role of hsps and related proteins in normal growth; the work discussed demonstrates a vital role for at least some of these proteins. The question whether the heat-inducible proteins perform the same function as the constitutively expressed proteins or carry out specialized functions remains unanswered. It is possible that the heat inducibility of some members of multigene families has evolved merely to increase the amount of total protein. However, it is reasonable to suppose that those proteins whose expression increases after stress, although perhaps able to perform some of the same functions as their noninducible relatives, have evolved the capacity to specifically cope with the physiological stresses rendered by heat and other

insults. Hopefully, over the next few years, functional distinctions between heat inducible and constitutive proteins and their roles in thermotolerance will be elucidated.

The diverse processes in which hsps function have been implicated indicates that these proteins are involved in many cellular processes. The job ahead for workers in the field is to determine which of these biochemical interactions observed in vitro, actually occur in vivo and which are essential for normal cell growth. The powerful approaches of both genetics and biochemistry will be needed to answer these questions.

ACKNOWLEDGMENTS:

We are grateful to many colleagues for sharing their unpublished results with us: R. Anderson, E. Baulieu, J. Bonner, I. Brown, P. Chambon, N. Collier, C. Georgopoulos, M. J. Gething, A. Goldberg, C. Gross, G. Hahn, R. Hallberg, L. Hightower, H. Iida, J. Landry, A. Lee, R. Morimoto, H. Pelham, D. Rose, J. Sambrook, M. Schlessinger, J. Silver, D. Toft, G. Walker, G. Weiderech, W. Welch, D. Young, R. Young. We thank R. Susek for help with figures.

Literature Cited

1. Adkins, B., Hunter, T., Sefton, B. M. 1982. The transforming proteins of PRCII virus and Rous sarcoma virus form a complex with the same two cellular phosphoproteins. *J. Virol.* 43:448–55
2. Allen, R., O'Brien, D., Eddy, E. 1988. A novel hsp70-like protein (P70) is present in mouse spermatogenic cells. *Mol. Cell. Biol.* 8:828–32
3. Anathan, J., Goldberg, A. L., Voellmy, R. 1986. Abnormal proteins serve as eucaryotic stress signals and trigger the activation of heat shock genes. *Science* 232:522–24
4. Anderson, R. L., Kersen, I. V., Basrur, V., Adwanker, M., Hahn, G. 1988. A new 70kD protein in heat resistant mammalian cells. *J. Cell Biochem. Suppl.* 12D p. 286
5. Ang, D., Chandrasekhar, G. N., Zylicz, M., Georgopoulos, C. 1986. *Escherichia coli grpE* gene codes for heat shock protein B25.3, essential for both lambda DNA replication at all temperatures and host growth at high temperature. *J. Bacteriol.* 167:25–29
6. Ardeshir, F., Flint, J. E., Richman, S. J., Reese, R. T. 1987. A 75 kd merozoite surface protein of *Plasmodium falciparum* which is related to the 70 kd heat-shock proteins. *EMBO J.* 6:493–99
7. Arrigo, A-P., Darlix, J. L., Khandjian, E. W., Simon, M., Spahr, P. F. 1985. Characterization of the prosome from Drosophila and its similarity to the cytoplasmic structures formed by the low molecular weight heat-shock proteins. *EMBO J* 4:399–406
8. Arrigo, A-P., Tanaka, K., Goldberg, A. L., Welch, W. J. 1988. Identity of the 19S 'prosome' particle with the large multifunctional protease complex of mammalian cells (the proteasome). *Nature* 31:192–94
9. Arrigo, A-P., Welch, W. J. 1987. Characterization and purification of the small 28000-dalton mammalian heat shock protein. *J. Biol. Chem.* 262: 15359–69
10. Ayme, A., Tissieres, A. 1985. Locus 67B of *Drosophila melanogaster* contains seven, not four, closely related heat shock genes. *EMBO J.* 4:2949–54
11. Bardwell, J. C. A., Craig, E. A. 1984. Major heat shock gene of *Drosophila* and the *Escherichia coli* heat inducible *dnaK* gene are homologous. *Proc. Natl. Acad. Sci. USA* 81:848–52
12. Bardwell, J. C. A., Craig, E. A. 1987. Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 84:5177–81
13. Bardwell, J. C. A., Craig, E. A. 1988. Ancient heat shock gene is dispensable. *J. Bacteriol.* 7:2977–83

14. Baulieu, E-E. 1987. Steroid hormone antagonists at the receptor level: a role for the heat-shock protein MW 90,000 (hsp 90). *J. Cell Biol.* 35:161–74
15. Berger, E. M., Woodward, M. P. 1983. Small heat shock proteins in *Drosophila* may confer thermal tolerance. *Exp. Cell Res.* 147:437–42
16. Bianco, A. E., Favaloro, J. M., Burkot, T. R., Culvenor, J. G., Crewther, P. E., et al. 1986. A repetitive antigen of *Plasmodium falciparum* that is homologous to heat shock protein 70 of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 83:8713–17
17. Binart, N., Chambraud, B., Dumas, B., Bigogne, C., Levin, J. M., et al. The cDNA-derived amino acid sequence of chicken HSP90 reveals a potential site of interaction with the DNA binding region of steroid receptors. Submitted
18. Blackman, R. K., Meselson, M. 1986. Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the Drosophila hsp82 gene. *J. Mol. Biol.* 188:499–515
19. Bloemendal, H., Berns, T., Zweers, A., Hoenders, H., Benedetti, E. L. 1972. The state of aggregation of a-crystallin detected after large-scale preparation by zonal centrifugation. *Eur. J. Biochem.* 24:401
20. Bochner, B., Zylicz, M., Georgopoulos, C. 1986. *Escherichia coli* DnaK protein possesses 5'-nucleotidase activity that is inhibited by AppppA. *J. Bacteriol.* 168:931–35
21. Bond, U., Schlesinger, M. J. 1985. Ubiquitin is a heat shock protein in chicken embryo fibroblasts. *Mol. Cell. Biol.* 5:949–56
22. Bond, U., Schlesinger, M. J. 1986. The chicken ubiquitin gene contains a heat shock promoter and expresses an unstable mRNA in heat-shocked cells. *Mol. Cell. Biol.* 6:4602–10
23. Braell, W. A., Schlossman, D. M., Schmid, S. L., Rothman, J. E. 1984. Dissociation of clathrin coats coupled to the hydrolysis of ATP: role of an uncoating ATPase. *J. Cell Biol.* 99:734–41
24. Bresnick, E. H., Redmond, T., Sanchez, E. R., Pratt, W. B., Welsh, M. J. 1988. Demonstration that the 90-kDa heat shock protein is associated with tubulin in L cell cytosol and in intact PtK cells. See Ref. 4, p. 283
25. Browder, L. W., Pollock, R. M., Nickells, R. W., Heikkila, J. J., Winning, R. S. 1988. Developmental regulation of the heat shock response. In *Genomic Adaptability in Cell Specialization*, ed. M. A. Bernardino, L. Etrin. New York: Plenum. In press
26. Brugge, J. S., Erikson, E., Erikson, R. L. 1981. The specific interaction of the Rous sarcoma virus transforming protein, $pp60^{src}$, with two cellular proteins. *Cell* 25:363–72
27. Brugge, J. S., Yuonemoto, W., Darrow, D. 1983. Interaction between the Rous sarcoma virus transforming protein and two cellular phosphoproteins: analysis of the turnover and distribution of this complex. *Mol. Cell. Biol.* 4:2697–2704
28. Brunt, S. A., Silver, J. C. 1986. Cellular localization of steroid hormone-regulated proteins during sexual development in *Achlya*. *Exp. Cell Res.* 165:306–19
29. Burshell, A., Stathis, P. A., Do, Y., Miller, S. C., Feldman, D. 1984. Characterization of an estrogen-binding protein in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* 259:3450–56
30. Cameron, S., Levin, L., Zoller, M., Wigler, M. 1988. cAMP-independent control of sporulation, glycogen metabolism, and heat shock resistance in *S. cerevisiae*. *Cell* 53:555–66
31. Carbajal, M. E., Duband, J. L., Lettre, F., Valet, J. P., Tanguay, R. M. 1986. Cellular localization of *Drosophila* 83-kilodalton heat shock protein in normal, heat-shocked, and recovering cultured cells with a specific antibody. *Biochem. Cell Biol.* 64:816–25
32. Carper, S. W., Duffy, J. J., Gerner, E. W. 1987. Heat-shock proteins in thermotolerance and other cellular process. *Cancer Res.* 47:5249–55
33. Catelli, M. G., Radanyi, C., Renoir, J. M., Binart, N., Baulieu, E. E. 1988. Definition of domain of hsp90 interacting with steroid receptors. See Ref. 4, p. 276
34. Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R., Georgopoulos, C. 1986. Purification and properties of the *groES* morphogenetic protein of *Escherichia coli*. *J. Biol. Chem.* 261: 12414–19
35. Chang, S. C., Wooden, S. K., Nakaki, T., Kim, Y. K., Lin, A. Y., et al. 1987. Rat gene encoding the 78-kDa glucose-regulated protein GRP78: its regulatory sequences and the effect of protein glycosylation on its expression. *Proc. Natl. Acad. Sci. USA* 84:680–84
36. Chappell, T. G., Welch, W. J., Schlossman, D. M., Palter, K. B., Schlesinger, M. J., Rothman, J. E. 1986. Uncoating ATPase is a member of the 70 kilodalton family of stress proteins. *Cell* 45:3–13

37. Cheney, C. M., Shearn, A. 1983. Developmental regulation of Drosophila disc proteins: synthesis of a heat shock protein under non-heat shock conditions. *Dev. Biol.* 95:325–30
38. Chiroco, W. J., Waters, M. G., Blobel, G. 1988. 70K heat shock related proteins stimulate protein translocation into microsomes. *Nature* 332:805–10
39. Chung, C. H., Goldberg, A. L. 1981. Studies of the ATP-dependent proteolytic enzyme protease La, from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 78:4728–32
40. Clarke, C. F., Cheng, K., Frey, A. B., Stein, R., Hinds, P. W., Levine, A. J. 1988. Purification of complexes of nuclear oncogene p53 with rat and *Escherichia coli* heat shock proteins: in vitro dissociation of hsc70 and DnaK from murine p53 by ATP. *Mol. Cell. Biol.* 8:1206–15
41. Collier, N. C., Heuser, J., Levy, M. A., Schlesinger, M. J. 1988. Ultrastructural and biochemical analysis of the stress granule in chicken embryo fibroblasts. *J. Cell Biol.* 106:1131–39
42. Collier, N. C., Schlesinger, M. J. 1986. The dynamic state of heat shock proteins in chicken embryo fibroblasts. *J. Cell Biol.* 103:1495–1507
43. Courtneidge, S. A., Bishop, J. M. 1982. Transit of $pp60^{v\text{-}src}$ to the plasma membrane. *Proc. Natl. Acad. Sci. USA* 79:7117–21
44. Craig, E. A., Ingolia, T. D., Manseau, L. J. 1983. Expression of Drosophila heat-shock cognate genes during heat shock and development. *Dev. Biol.* 99:418–26
45. Craig, E. A., Jacobsen, K. 1984. Mutations of the heat-inducible 70 kilodalton genes of yeast confer temperature sensitive growth. *Cell* 38:841–49
46. Craig, E. A., Jacobsen, K. 1985. Mutations in cognate genes of *Saccharomyces cerevisiae* hsp70 result in reduced growth rate at low temperatures. *Mol. Cell. Biol.* 5:3517–24
47. Craig, E. A., Kramer, J., Kosic-Smithers, J. 1987. *SSC1*, a member of the 70-kDa heat shock protein multigene family of *Saccharomyces cerevisiae*, is essential for growth. *Proc. Natl. Acad. Sci. USA* 84:4156–60
48. Craig, E., Slater, M., Stone, D., Park, H. O., Boostein, W. 1987. Regulation of a yeast heat shock gene. In RNA polymerase and the regulation of transcription, ed. W. Reznikoff, R. Burgess, J. Dahlberg, C. Gross, T. T. Record, M. Wickens, Elsevier Sci. pp. 267–278
49. Deleted in text
50. Czarnecka, E., Gurley, W. B., Nagao, R. T., Mosquera, L. A., Key, J. L. 1985. DNA sequence and transcript mapping of a soybean gene encoding a small heat shock protein. *Proc. Natl. Acad. Sci. USA* 82:3726–30
51. Daniels, C. J., McKee, A. H. Z., Doolittle, W. F. 1984. Archaebacterial heat shock proteins. *EMBO J.* 3:745–49
52. Daskal, Y., Smetana, K., Busch, H. 1980. Evidence from studies on segregated nucleoli that nucleolar silver staining proteins C23 and B23 are in the fibrillar component. *Exp. Cell Res.* 127:285–91
53. Deshaies, R., Koch, B., Werner-Washburne, M., Craig, E., Scheckman, R. 1988. 70kd stress protein homologues facilitate translocation of secretory and mitochondrial precursor polypeptides. *Nature* 332:800–05
54. Dorner, A., Krane, M., Kaufman, R. 1988. Reduction of endogenous grp 78 levels in CHO cells and its effect on secretion. *J. Cell. Biochem. Suppl.* 12D:276
55. Dougherty, J. J., Puri, R. K., Toft, D. O. 1984. Polypeptide components of two 8 S forms of chicken oviduct progesterone receptor. *J. Biol. Chem.* 259:8004–09
56. Dragon, E. A., Sias, S. R., Kato, E. A., Gabe, J. D. 1987. The genome of Trypanosoma cruzi contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. *Mol. Cell Biol.* 7:1271–75
57. Duncan, R., Hershey, J. W. 1984. Heat shock-induced translational alterations in HeLa cells. *J. Biol. Chem.* 259: 11882–89
58. Echols, H. 1986. Multiple DNA-protein interactions governing high-precision DNA transactions. *Science* 233:1050–56
59. Eilers, M., Oppliger, W., Schatz, G. 1987. Both ATP and an energized inner membrane are required to import a purified precursor protein into mitochondria. *EMBO J.* 6:1073–77
60. Eissenberg, J. C., Elgin, S. C. R. 1987. $Hsp28^{ah}$: A *P*-element insertion mutation that alters the expression of a heat shock gene in *Drosophila melanogaster*. *Genetics* 115:333–40
61. Ellis, J. 1987. Proteins as molecular chaperones. *Nature* 328:378–79
62. Ellwood, M. S., Craig, E. A. 1984. Differential regulation of the 70K heat shock and related genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 4:1454–59
63. Evans, R. M. 1988. The steroid and

thyroid hormone receptor superfamily. *Science* 240:889–95
64. Falkenburg, P. E., Haass, C., Kloetzel, P. M., Niedel, B., Kopp, F., et al. 1988. Drosophila small cytoplasmic 19S ribonucleoprotein is homologous to the rat multicatalytic proteinase. *Nature* 331:190–92
65. Farrelly, F. W., Finkelstein, D. B. 1984. Complete sequence of the heat shock-inducible *HSP90* gene of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 9:5745–51
66. Fayet, O., Louarn, J., Georgopoulos, C. 1986. Suppression of the *Escherichia coli dnaA46* mutation by amplification of the *groES* and *groEL* genes. *Mol. Gen. Genet.* 202:435–45
67. Ferrini, U., Falcioni, R., Delpino, A., Cavaliere, R., Zupi, G., Natali, P. G. 1984. The heat-shock proteins produced by two human myeloma cell lines: absence of correlation with thermosensitivity. *Int. J. Cancer* 34:651–55
68. Ferris, D. K., Harel-Bellan, A., Morimoto, R. I., Welch, W. J., Farrar, W. L. Mitogen and kymphokine stimulation of heat shock proteins in T lymphocytes. Submitted to *Proc. Natl. Acad. Sci. USA*
69. Findly, R. C., Gilies, R. J., Shulman, R. G. 1983. In vivo Phosphorus-31 nuclear magnetic resonance reveals lowered ATP during heat shock of Tetrahymena. *Science* 219:1223–25
70. Finlay, C. A., Hinds, P. W., Tan, T. H., Eliyahu, D., Oren, M., Levine, A. J. 1988. Activating mutations for transformation by p53 produce a gene product that forms an hsc70-p53 complex with an altered half-life. *Mol. Cell. Biol.* 8:531–39
71. Finley, D., Crechanover, A., Varshavsky, A. 1984. Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* 37:43–55
72. Finley, D., Ozkaynak, E., Varshavsky, A. 1987. The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation and other stresses. *Cell* 48:1035–46
73. Georgopoulos, C. P. 1977. A new bacterial gene *(groPC)* which affects lambda DNA replication. *Mol. Gen. Genet.* 151:35–39
74. Georgopoulos, C. P., Hendrix, R. W., Casjens, S. R., Kaiser, A. D. 1973. Host participation in bacteriophage lambda head assembly. *J. Mol. Biol.* 76: 45–60
75. Gething, M. J., McCammon, K., Sambrook, J. 1986. Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport. *Cell* 46:939–50
76. Giebel, L. B., Dworniczak, B. P., Bautz, E. K. F. 1988. Developmental regulation of a constitutively expressed mouse mRNA encoding a 72-kDa heat shock-like protein. *Dev. Biol.* 125:200–207
77. Gilmour, D. S., Lis, J. 1986. RNA polymerase II interacts with the promoter region of the noninduced *hsp70* gene in *Drosophila melanogaster* cells. *Mol. Cell Biol.* 6:3984–89
78. Glaser, R. L., Wolfner, M. F., Lis, J. T. 1986. Spatial and temporal pattern of HSP26 expression during normal development. *EMBO J.* 5:747–54
79. Goate, A. M., Cooper, D. N., Hall, C., Leung, T. K. C., Solomon, E., Lim, L. 1987. Localization of a human heat-shock HSP70 gene sequence to chromosome 6 and detection of two other loci by somatic-cell hybrid and restriction fragment length polymorphism analysis. *Hum. Genet.* 75:123–28
80. Goff, S. A., Goldberg, A. L. 1987. An increased content of protease La, the *lon* gene product, increases protein degradation and blocks growth in *Escherichia coli*. *J. Biol. Chem.* 262:4508–15
81. Greene, L., Eisinberg, E. 1988. The dissociation of clathrin from coated vesicles by the uncoating ATPase. *J. Cell. Biochem. Suppl.* 12D:278
82. Grossman, A. D., Erickson, J. W., Gross, C. A. 1984. The htpR gene of *Escherichia coli* is a sigma factor for heat-shock promoters. *Cell* 28:282–90
83. Groyer, A., Schweizer-Groyer, G., Cadepond, F., Mariller, M., Baulieu, E-E. 1987. Antiglucocorticosteroid effects suggest why steroid hormone is required for receptors to bind DNA in vivo but not in vitro. *Nature* 328:624–26
84. Guest, J. R., Nice, H. M. 1978. Chromosomal location of the *mop (groE)* gene necessary for bacteriophage morphogenesis in *Escherichia coli*. *J. Gen. Microbiol.* 109:329–33
85. Guidon, P. T. Jr., Hightower, L. 1985. The major heat shock cognate protein of rat liver is a lipoprotein. *J. Cell. Biol.* 99(Pt. 2):454a
86. Guidon, P. T., Hightower, L. E. 1986. Purification and initial characterization of the 71-kilodalton rat heat-shock protein and its cognate as fatty acid binding proteins. *Biochemistry* 25:3231–39
87. Haas, I. G., Wabl, M. 1983. Immunoglobulin heavy chain binding protein. *Nature* 306:387–89
88. Harrison, G. S., Drabkin, H. A., Kao, F., Hartz, J., Hart, I. M., et al. 1987. Chromosomal location of human genes encoding major heat-shock protein HSP70. *Somatic Cell Mol. Genet.* 13: 119–30
89. Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., et al. 1988. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature.* 333:330–334
90. Hendrix, R. W. 1979. Purification and properties of GroE, a host protein involved in bacteriophage assembly. *J. Mol. Biol.* 129:375–92
91. Hickey, E., Brandon, S. E., Potter, R., Stein, G., Stein, J., Weber, L. A. 1986. Sequence and organization of genes encoding the human 27 kDa heat shock protein. *Nucleic Acids Res.* 14:4127–45
92. Hinds, P. W., Finlay, C. A., Frey, A. B., Levine, A. J. 1987. Immunological evidence for the association of p53 with a heat shock protein, hsc70, in p53-plus-ras-transformed cell lines. *Mol. Cell. Biol.* 7:2863–69
93. Horiuchi, T., Maki, H., Sekiguchi, M. 1984. RNase H-defective mutants of *Escherichia coli*: a possible discriminatory role of RNase H in initiation of DNA replication. *Mol. Gen. Genet.* 195:17–22
94. Iida, H., Yahara, I., 1985. Yeast heat-shock protein of Mr 48,000 is an isoprotein of enolase. *Nature* 315:688–90
95. Ingolia, T. D., Craig, E. A. 1982. Drosophila gene related to the major heat-shock-induced gene is transcribed at normal temperatures and not induced by heat shock. *Proc. Natl. Acad. Sci. USA* 79:525–29
96. Ingolia, T. D., Craig, E. A. 1982. Four small Drosophila heat shock proteins are related to each other and to mammalian a-crystallin. *Proc. Natl. Acad. Sci. USA* 79:2360–64
97. Jenkins, A. J., March, J. B., Oliver, I. R., Musters, M. 1986. A DNA fragment containing the *groE* genes can suppress mutations in the *Escherichia coli dnaA* gene. *Mol. Gen. Genet.* 202:446–54
98. Joab, I., Radanyi, C., Renoir, M., Buchou, T., Catelli, M-G., et al. 1984. Common non-hormone binding component in non-transformed chick oviduct receptors of four steroid hormones. *Nature* 308:850–53
99. Jones, K. A., Findly, R. C. 1986. Induction of HSPs by canavanine in *Tetrahymena*. *J. Biol. Chem.* 261:8703–07
100. Joplin, K. H., Chen, C-P., Yocum, G. D., Denlinger, D. L. 1988. Induction of a tissue specific set of heat shock proteins in the flesh fly *sarcophaga crassipalpis*. See Ref. 4, p. 270
101. Keene, M. A., Corces, V., Lowenhaupt, K., Elgin, S. C. R. 1981. DNase I hypersensitive sites in *Drosophila* chromatin at the 5' ends of regions of transcription. *Proc. Natl. Acad. Sci. USA* 78:143–46
102. Ketis, N.V., Lawler, J., Karnovsky, M. J. 1988. Heat shock stimulates the expression of thrombospondin by endothelial cells in culture. See Ref. 4, p. 261
103. Kim, Y-J., Shuman, J., Sette, M., Przybyla, A. 1984. Nuclear localization and phosphorylation of three 25-kilodalton rat stress proteins. *Mol. Cell. Biol.* 4:468–74
104. Kimpel, J. A., Key, J. L. 1985. Presence of heat shock mRNAs in field grown soybeans. *Plant Physiol.* 79:672–78
105. Kistler, J., Dumcombe, Y., Laimmli, U. K. 1984. Mapping nucleolar proteins with monoclonal antibodies. *J. Cell. Biol.* 99:1981–88
106. Kloetzel, P., Bautz, E. K. F. 1983. Heat-shock proteins are associated with hnRNP in *Drosophila melanogaster* tissue culture cells. *EMBO J.* 2:705–10
107. Kochan, J., Murialdo, H. 1983. Early intermediates in bacteriophage lambda prohead assembly. *Virology* 131:100–15
108. Koyasu, S., Nishida, E., Kadowaki, T., Matsuzaki, F., Iida, K., et al. 1986. Two mammalian heat shock proteins, HSP90 and HSP100 are actin-binding proteins. *Proc. Natl. Acad. Sci. USA* 83:8054–58
109. Kulomaa, M. S., Weigel, N. L., Kleinsek, D. A., Beattie, W. G., Conneely, O. M., et al. 1986. Amino acid sequence of a chicken heat shock protein derived from the complementary DNA nucleotide sequence. *Biochemistry* 25:6244–51
110. Kurtz, S., Lindquist, S. L. 1984. The changing pattern of gene expression in sporulating yeast. *Proc. Natl. Acad. Sci. USA* 81:7323–27
111. Kurtz, S., Rossi, J., Petko, L., Lindquist, S. 1986. An ancient development induction: heat-shock proteins induced in sporulation and oogenesis. *Science* 231:1154–57
112. Kyte, J., Doolittle, R. F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105–32
113. Lai, B. T., Chin, N. W., Stanek, A. E., Keh, W., Lanks, K. W. 1984. Quantitation and intracellular localization of the 85K heat shock protein by using mono-

clonal and polyclonal antibodies. *Mol. Cell. Biol.* 4:2802–10
114. Landick, R., Vaughn, V., Lau, E. T., VanBogelen, R. A., Erickson, J. W., Neidhardt, F. C. 1984. Nucleotide sequence of the heat shock regulatory gene of *Escherichia coli* suggests its protein product may be a transcription factor. *Cell* 38:175–82
115. Landry, J. Temperature resistant variants of CHO synthesize hsp28 constitutively. See Ref. 4, p.212
116. Latchman, D. S., Chan, W. L., Leaver, C. E., Patel, R., Oliver, P., LaThangue, N. B. 1987. The human Mr 90,000 heat shock protein and the *Escherichia coli* Lon protein share an antigenic determinant. *Comp. Biochem. Physiol.* 87:961–67
117. LeBowitz, J. H., McMacken, R. 1986. The Escherichia coli *dnaB* replication protein is a DNA helicase. *J. Biol. Chem.* 261:4738–48
118. LeBowitz, J., Roberts, J. D., McMacken, R. 1986. Specialized nucleoprotein structures at the origin of replication of bacteriophage lambda: Localized unwinding of duplex DNA by a six protein reaction. *Proc. Natl. Acad. Sci. USA* 83:7638–42
119. Lee, A. S., Bell, J., Ting, J. 1984. Biochemical characterization of the 94- and 78-kilodalton glucose-regulated proteins in hamster fibroblasts. *J. Biol. Chem.* 259:4616–21
120. Lee, Y. J., Dewey, W. C. 1988. Thermotolerance induced by heat, sodium aresenite, or puromycin: its inhibition and differences between 43°C and 45°C. *J. Cell Physiol.* 136:991–
121. Lewis, M. J., Pelham, H. R. B. 1985. Involvement of ATP in the nuclear and nucleolar functions of the 70 kd heat shock protein. *EMBO J.* 4:3137–43
122. Li, G. C., Laszlo, A. 1985. Thermotolerance in mammalian cells: a possible role for the heat shock proteins. *Changes in Eukaryotic Gene Expression in Response to Environmental Stress*, pp. 349–71. London: Academic
123. Lin, J. J. C., Welch, W. J., Garrels, J. I., Feramisco, J. R. 1982. The association of the 100kd heat shock protein with the Golgi apparatus. In *Heat Shock: From Bacteria to Man*, ed. M. J. Schlesinger, M. Ashburner, A. Tissieres, pp. 267–73. New York: Cold Spring Harbor Lab.
124. Lindahl, G., Lindahl, T. 1984. Initiation of DNA replication in *Escherichia coli*: RNase H-defective mutants do not require the *dnaA* function. *Mol. Gen. Genet.* 196:283–89
125. Lindquist, S. 1980. Varying patterns of protein synthesis during heat shock: implications for regulation. *Dev. Biol.* 77:463–79
126. Lindquist, S. 1981. Regulation of protein synthesis during heat shock. *Nature* 293:311–14
127. Lindquist, S. 1986. The heat-shock response. *Ann. Rev. Biochem.* 55:1151–91
128. Lipsich, L. A., Cutt, J. R., Brugge, J. S. 1982. Association of the transforming proteins of Rous, Fujinami, and Y73 avian sarcoma viruses with the same two cellular proteins. *Mol. Cell. Biol.* 2:875–80
129. Loomis, W., Wheeler, S., Schmidt, S. 1982. Phosphorylation of the major heat shock protein of *Dictyostelium discoideum*. *Mol. Cell. Biol.* 2:484–89
129a. Loomis, W. F., Wheeler, S. A. 1982. Chromatin-associated heat shock proteins of Dictyostelium. *Dev. Bio.* 90: 412–18
130. Lowe, D. G., Moran, L. A. 1986. Molecular-cloning and analysis of DNA complementary to three mouse Mr=68,000 heat-shock protein mRNAs. *J. Biol. Chem.* 261:2102–12
131. Mansfield, M. A., Key, J. L. 1987. Synthesis of the low molecular weight heat shock proteins in plants. *Plant Physiol.* 84:1007–17
132. Mariano, T. M., Siekierka, J. 1986. Inhibition of HeLa cell protein synthesis under heat shock conditions in the absence of initiation factor eIF-2 alpha phosphorylation. *Biochem. Biophys. Res. Commun.* 138:519–25
133. Mazzarella, R. A., Green, M. 1987. ERp99, an abundant, conserved glycoprotein of the endoplasmic reticulum, is homologous to the 90-kDa heat shock protein (hsp90) and the 94-kDa glucose regulated protein (GRP94). *J. Biol. Chem.* 262:8875–83
134. McAlister, L., Finkelstein, D. B. 1980. Heat shock proteins and thermal resistance in yeast. *Biochem. Biophys. Res. Commun.* 93:819–24
135. McGarry, T. J., Lindquist, S. 1986. Inhibition of heat-shock protein-synthesis by heat-inducible antisense RNA. *Proc. Natl. Acad. Sci. USA* 83:399–404
136. McMullin, T. W., Hallberg, R. L. 1987. A normal mitochondrial protein is selectively synthesized and accumulated during heat shock in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 7:4414–23
137. McMullin, T. W., Hallberg, R. L. 1988. A highly evolutionary conserved mitochondrial protein is structurally related to the protein encoded by the *Escherichia coli groEL* gene. *Mol. Cell. Biol.* 8:371–80

138. Mendel, D. B., Orti, E. 1988. Isoform composition and stoichiometry of the 90 kDa heat shock protein associated with glucocorticoid receptors. *J. Biol. Chem.* 263:6695–702
139. Metzger, D., White, J. H., Chambon, P. 1988. The human oestrogen receptor functions in yeast. *Nature*. 334:31–36
140. Milarski, K. L., Morimoto, R. 1986. Expression of human HSP70 during the synthetic phase of the cell cycle. *Proc. Natl. Acad. Sci. USA* 83:9517–21
141. Mitchell, H., Moller, G., Peterson, N., Lipps-Sarmiento, L. 1979. Specific protection from phenocopy induction by heat shock. *Dev. Genet.* 1:191–93
142. Mizusawa, S., Gottesman, S. 1983. Protein degradation in *Escherichia coli*: the *lon* gene controls the stability of the *sulA* protein. *Proc. Natl. Acad. Sci. USA* 80:358–62
143. Moore, S. K., Kozak, C., Robinson, E. A., Ullrich, S. J., Appella, E. 1987. Cloning and nucleotide sequence of the murine hsp84 cDNA and chromosome assignment of related sequences. *Gene* 56:29–40
144. Mues, G. I., Munn, T. Z., Raese, J. D. 1986. A human gene family with sequence homology to *Drosophila melanogaster* Hsp70 heat shock gene. *J. Biol. Chem.* 26:874–77
145. Munro, S., Pelham, H. 1985. What turns on heat-shock genes? *Nature* 317:477–78
146. Munro, S., Pelham, H. 1986. An hsp70-like protein in the ER: identity with the 78kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* 46:291–300
147. Munro, S., Pelham, H. R. B. 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48:899–907
148. Nagao, R. T., Kimpel, J. A., Vierling, E., Key, J. L. 1986. The heat shock response: a comparative analysis. In *Oxford Surveys of Plant Molecular and Cell Biology*, ed. B. J. Miflin, p. 384. Oxford: Oxford Univ. Press
149. Neidhardt, F., VanBogelen, R. 1987. "Heat shock response." In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. F. C. Neidhardt, pp. 1334–45. Washington, DC: Am. Soc. Microbiol.
150. Nene, V., Dunne, D. W., Johnson, K. W., Taylor, D. W., Cordingley, J. S. 1986. Sequence and expression of a major egg antigen from *Schistosoma mansoni*. Homologies to heat shock proteins and alpha-crystallins. *Mol. Biochem. Parasitol.* 21:179–88
151. Nerland, A. H., Mustafa, A. S., Sweetser, D., Godal, T., Young, R. 1988. A protein antigen of *Mycobacterium leprea* is related to a family of small heat shock proteins. Manuscript submitted
152. Neumann, D., zur Nieden, U., Mateuffel, R., Walter, G., Scharf, K-D., Nover, L. 1987. Intracellular localization of heat-shock proteins in tomato cell cultures. *Eur. J. Cell Biol.* 43:71–81
153. Nickells, R. W. 1987. The developmental acquisition and mechanism of thermotolerance in *Xenopus laevis* embryos. PhD thesis. Dep. Biol. Sci., Calgary, Alberta, Canada
154. Nishida, E., Koyasu, S., Sakai, H., Yahara, I. 1986. Calmodulin-regulated binding of the 90-kDa heat shock protein to actin filaments. *J. Biol. Chem.* 261:16033–36
155. Nover, L., Munsche, D., Neumann, D., Ohme, K., Scharf, K. D. 1986. Control of ribosome biosynthesis in plant cell cultures under heat-shock conditions. Ribosomal RNA. *Eur. J. Biochem.* 160:297–304
156. Nover, L., Scharf, K. 1984. Synthesis, modification and structural binding of heat-shock proteins in tomato cells. *Eur. J. Biochem.* 139:303–08
157. Nover, L., Scharf, K. D., Neumann, D. 1988. Cytoplasmic heat shock granules are formed from precursor particles and contain a specific set of mRNAs. Submitted to *Mol. Cell. Biol.*
158. Oppermann, H., Levinson, W., Bishop, J. M. 1981. A cellular protein that associates with the transforming protein of Rous sarcoma virus is also a heat-shock protein. *Proc. Natl. Acad. Sci. USA* 78:1067–71
159. Ozkaynak, E., Finley, D., Solomon, M., Varshavsky, A. 1987. The yeast ubiquitin genes: a family of natural gene fusions. *EMBO J.* 6:1429–39
160. Ozkaynak, E., Finley, D., Varshavsky, A. 1984. The yeast ubiquitin gene: head-to-tail repeats encoding a polyubiquitin precursor protein. *Nature* 312:663–66
161. Paek, K., Walker, G. C. 1987. *Escherichia coli dnaK* mutants are inviable at high temperature. *J. Bacteriol.* 169:283–90
162. Palter, K. B., Watanabe, M., Stinson, L., Mahowald, A. P., Craig, E. A. 1986. Expression and localization of *Drosophila melanogaster* hsp70 cognate proteins. *Mol. Cell. Biol.* 6:1187–1203
163. Pauli, D., Tonka, C-H., Ayme-Southgate, A. 1988. An unusual split *Dro-*

sophila heat shock gene expressed during embryogenesis, pupation and in testis. *J. Mol. Biol.* 200:47–53
164. Pedersen, R. C., Brownie, A. C. 1987. Steroidogenesis-activator polypeptide isolated from a rat leydig cell tumor. *Science* 236:188–90
165. Pelham, H. R. B. 1984. HSP70 accelerates the recovery of nucleolar morphology after heat shock. *EMBO J.* 3:3095–100
166. Pelham, H. R. B. 1986. Speculations on the functions of the major heat shock and glucose-regulated proteins. *Cell* 46:959–61
167. Petersen, R., Linquist, S. 1988. The *Drosophila hsp70* message is rapidly degraded at normal temperatures and stabilized by heat shock. *Gene*. In press
168. Petko, L., Lindquist, S. 1986. Hsp26 is not required for growth at high temperatures, nor for thermotolerance, spore development, or germination. *Cell* 45: 885–94
169. Petko, L., McGarry, T., Taulien, J., Lindquist, S. 1988. Hsp35 is a member of the glyeraldehyde-3-phosphate dehydrogenase gene family. Manuscript submitted
170. Pfanner, N., Neupert, W. 1986. Transport of F1-ATPase subunit B into mitochondria depends on both a membrane potential and nucleoside triphosphates. *FEBS Lett.* 209:152–56
171. Pinhasi-Kimhi, O., Michalovitz, D., Ben-Zeev, A., Oren, M. 1986. Specific interaction between the p53 cellular tumour antigen and major heat shock proteins. *Nature* 320:182–84
172. Piper, P. W., Curran, B., Davies, M. W., Lockherat, A., Feid, G. 1986. Transcription of the phosphoglycerate kinase one of Saccharomyces cerevisiae increases when fermentative cultures are stressed by heat shock. *Eur. J. Biochem.* 161:525–31
173. Polaina, J., Conde, J. 1982. Genes involved in the control of nuclear fusion during the sexual cycle of *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* 186:253–58
174. Redeuilh, G., Moncharmont, B., Secco, C., Baulieu, E-E. 1987. Subunit composition of the molybdate-stabilized "8-9 S" nontransformed estradiol receptor purified from calf uterus. *J. Biol. Chem.* 262:6969–75
175. Renoir, J-M., Buchou, T., Baulieu, E-E. 1986. Involvement of a nonhormone-binding 90-kilodalton protein in the nontransformed 8S form of the rabbit uterus progesterone receptor. *Biochemistry* 25:6405–13
176. Rimland, J., Akhayat, O., Infante, D., Infante, A. A. 1988. Developmental regulation and biochemical analysis of a 21 Kd heat shock in sea urchins. See Ref. 4, p. 271
177. Deleted in text
178. Rollet, E., Best-Belpomme, M. 1986. HSP 26 and 27 are phosphorylated in response to heat shock and ecdysterone in *Drosophila melanogaster* cells. *Biochem. Biophys. Res. Commun.* 141: 426–33
179. Rose, D. W., Wettenhall, R. E. H., Kudlicki, W., Kramer, G., Hardesty, B. 1987. The 90-kilodalton peptide of the heme-regulated eIF-2a kinase has sequence similarity with the 90-kilodalton heat shock protein. *Biochemistry* 26: 6583–87
180. Rossi, J., Lindquist, S. 1988. Intracellular location of hsp26 in yeast cells varies with metabolism. *J. Cell Biol.* In press
181. Rotii-Roti, J. L., Laszlo, A. 1987. The effects of hyperthermia on cellular macromolecules. In *Hyperthermia and Oncology*, Vol. 1. ed. M. Urano, A. Douple. The Netherlands, VNU Scientific Publishers. pp. 13–56
182. Russnak, R. H., Candido, P. M. 1985. Locus encoding a family of small heat shock genes in Caenorhabditis elegans: two genes duplicated to form a 34.8-kilobase inverted repeat. *Mol. Cell. Biol.* 5:1268–78
183. Saga, S., Nagata, K., Chen, W-T., Yamada, K. M. 1987. pH-dependent function, purification and intracellular location of a major collagen-binding glycoprotein. *J. Cell. Biol.* 105:517–27
184. Sakakibara, Y. 1988. The *dnaK* gene Escherichia coli functions in initiation of chromosome replication. *J. Bacteriol.* 170:972–79
185. Sanchez, E. R., Housley, P. R., Pratt, W. B. 1986. The molybdate-stabilized glucocorticoid binding complex of L-cells contains a 98–100 Kdalton steroid binding phosphoprotein and a 90 Kdalton nonsteroid-binding phosphoprotein that is part of the murine heat-shock complex. *J. Steroid Biochem.* 24:9–18
186. Sanchez, E. R., Meshinchi, S., Tienrungroj, W., Schlesinger, M. J., Toft, D. O., Pratt, W. B. 1987. Relationship of the 90-kDa murine heat shock protein to the untransformed and transformed states of the L cell glucocorticoid receptor. *J. Biol. Chem.* 262:6986–91
187. Schlossman, D. M., Schmid, S. L., Braell, W. A., Rothman, J. E. 1984. An enzyme that removes clathrin coats: purification of an uncoating ATPase. *J. Cell Biol.* 99:723–33

188. Schmid, H. P., Akhayat, O., Martins de Sa, C., Puvion, F., Koehler, K., Scherrer, K. 1984. The prosome: An ubiquitous morphologically distinct RNP particle associated with repressed mRNPs and containing specific scRNA and a characteristic set of proteins. *EMBO J.* 3:29–34
189. Schmid, S. L., Rothman, J. E. 1985. Enzymatic dissociation of clathrin cages in a two-stage process. *J. Biol. Chem.* 260:10044–49
190. Schmid, S. L., Rothman, J. E. 1985. Two classes of binding sites for uncoating protein in clathrin triskelions. *J. Biol. Chem.* 260:10050–56
191. Schuldt, C., Kloetzel, B. M. 1985. Analysis of cytoplasmic 19S ring-type particles in Drosophila which contain hsp23 at normal growth temperature. *Dev. Biol.* 11:65–74
192. Sciandra, J. J., Subjeck, J. R. 1983. The effects of glucose on protein synthesis and thermosensitivity in Chinese hamster ovary cells. *J. Biol. Chem.* 258: 12091–93
193. Shiu, R. P. C., Pouyssegur, J., Pastan, I. 1977. Glucose depletion accounts for the induction of two transformation-sensitive membrane proteins in Rous sarcoma virus-transformed chick embryo fibroblasts. *Proc. Natl. Acad. Sci. USA* 74:3840–44
194. Shyamala, G., Ramachandran, C., Catelli, M. G., Schneider, W. 1988. Estrogenic regulation of murine uterine 90 Kd heat shock protein. See Ref. 4, p. 292
195. Shyy, T-T., Subjeck, J. R., Heinaman, R., Anderson, G. 1986. Effect of growth state and heat shock on nucleolar localization of the 110,000-Da heat shock protein in mouse embryo fibroblasts. *Cancer Res.* 46:4738–45
196. Silver, J. C., Riehl, R., Brunt, S. 1988. Accumulation of mRNA for *Achlya* hsp85, a component of the *Achlya* steroid receptor can be induced by either heat shock or steroid hormone. *J. Cell Biochem.* 12:215–72
197. Sirotkin, K., Bartley, N., Perry, W. L., Briggs, D., Grell, E. H., et al. 1986. Deletion polymorphism in a *Drosophila melanogaster* heat shock gene. *Mol. Gen. Genet.* 204:266–72
198. Sirotkin, K., Davidson, N. 1982. Developmentally regulated transcription from *Drosophila melanogaster* chromosomal site 67B. *Dev. Biol.* 89:196–210
199. Sorger, P. K., Lewis, M. J., Pelham, H. R. B. 1987. Heat shock factor is regulated differently in yeast and HeLa cells. *Nature* 329:81–84
200. Sorger, P. K., Pelham, H. R. B. 1987. The glucose-regulated protein grp94 is related to heat shock protein hsp90. *J. Mol. Biol.* 194:341–44
201. Southgate, R., Ayme, A., Voellmy, R. 1983. Nucleotide sequence analysis of the *Drosophila* small heat shock gene cluster at locus 67B. *J. Mol. Biol.* 165:35–57
202. Storti, R. V., Scott, M. P., Rich, A., Pardue, M. L. 1980. Translational control of protein synthesis in response to heat shock in *D. melanogaster* cells. *Cell* 22:825–34
203. Straus, D. B., Walter, W. A., Gross, C. A. 1987. The heat shock response of *E. coli* is regulated by changes in the concentration of 32. *Nature* 329:348–51
204. Subjeck, J. R., Shyy, T., Shen, J., Johnson, R. J. 1983. Association between the mammalian 110,000-dalton heat-shock protein and nucleoli. *J. Cell Biol.* 97:1389–95
205. Subjeck, J. R., Shyy, T-T. 1986. Stress protein systems of mammalian cells. *Am. J. Physiol.* 250:C1-C17
206. Sunshine, M., Feiss, M., Stuart, J., Yochem, J. 1977. A new host gene *(groPC)* necessary for lambda DNA replication. *Mol. Gen. Genet.* 151:27–34
207. Theodorakis, N. G., Morimoto, R. I. 1987. Posttranscriptional regulation of hsp70 expression in human cells: effects of heat shock, inhibition of protein synthesis, and adenovirus infection on translation and mRNA stability. *Mol. Cell. Biol.* 7:4357–68
208. Thomas, S. R., Lengyel, J. A. 1986. Ecdysteroid-regulated heat-shock gene expression during *Drosophila melanogaster* development. *J. Cell. Physiol.* 127:451–56
209. Tilly, K., Erickson, J., Sharma, S., Georgopoulos, C. 1986. Heat shock regulatory gene rpoH mRNA level increases after heat shock in *Escherichia coli*. *J. Bacteriol.* 168:1155–58
210. Tilly, K., Georgopoulos, C. 1982. Evidence that the two *Escherichia coli groE* morphogenetic gene products interact in vivo. *J. Bacteriol.* 149:1082–1088
211. Tilly, K., Murialdo, H., Georgopoulos, C. P. 1981. Identification of a second *Escherichia coli groE* gene whose product is necessary for bacteriophage morphogenesis. *Proc. Natl. Acad. Sci. USA* 78:1629–33
212. Tilly, K., Van Bogelen, R. A., Georgopoulos, C., Neidhardt, F. C. 1983.

Identification of the heat-inducible protein C15.4 as the *groES* gene product in *Escherichia coli*. *J. Bacteriol.* 154: 1505–07
213. Tobe, T., Ito, K., Yura, T. 1984. Isolation and physical mapping of temperature-sensitive mutants defective in heat-shock induction of proteins in *Escherichia coli*. *Mol. Gen. Genet.* 195: 10–16
214. Tsuchido, T., VanBogelen, R. A., Neidhardt, F. C. 1986. Heat shock response in *Escherichia coli* influences cell division. *Proc. Natl. Acad. Sci. USA* 83:6959–63
215. Ullrich, S. J., Robinson, E. A., Law, L. W., Willingham, M., Appella, E. 1986. A mouse tumor-specific transplantation antigen is a heat shock-related protein. *Proc. Natl. Acad. Sci. USA* 83:3121–25
216. Ungewickell, E. 1985. The 70-kd mammalian heat shock proteins are structurally and functionally related to the uncoating protein that releases clathrin triskelion from coated vesicles. *EMBO J.* 4:3385–91
217. Van der Ouderaa, F. J., de Jong, W. W., Bloemendal, H. 1973. Amino-acid sequence alpha-A_2 chain of bovine alpha-crystallin. *Eur. J. Biochem.* 39: 207–22
218. van Bergen en Henegouwen, P. M. P., Berbers, G., Linnemans, W. A. M., vanWijk, R. 1987. Subcellular localization of the 84,000 dalton heat-shock protein in mouse neuroblastoma cells: evidence for a cytoplasmic and nuclear location. *Eur. J. Cell Biol.* 43:469–78
219. VanBogelen, R. A., Acton, M. A., Neidhardt, F. C. 1987. Induction of the heat shock regulon does not produce thermotolerance in *Escherichia coli*. *Genes Dev.* 1:525–31
220. Deleted in text
221. Vass, K., Welch, W. J., Nowak, T. S. Jr. 1988. Localization of 70 KDa stress protein induction in gerbil brain after ischemia. *Acta Neuropathol.* In press
222. Deleted in text
223. Velazquez, J. M., Lindquist, S. 1984. HSP70: nuclear function during environmental stress; cytoplasmic storage during recovery. *Cell* 36:655–62
224. Deleted in text
225. Verner, K., Schatz, G. 1987. Import of an incompletely folded precursor protein into isolated mitochondria requires an energized inner membrane, but no added ATP. *EMBO J.* 6:2449–56
226. Vierling, E., Nagao, R. T., DeRocher, A. E., Harris, L. M. 1988. A heat shock protein localized to chloroplasts is a member of a eukaryotic superfamily of heat shock proteins. *EMBO J.* 7:575–81
227. Wada, M., Fujita, H., Itikawa, H. 1987. Genetic suppression of a temperature-sensitive *groES* mutation by an altered subunit of RNA polymerase of *Escherichia coli* K-12. *J. Bacteriol.* 169:1102–06
228. Wada, M., Itikawa, H. 1984. Participation of *Escherichia coli* K-12 *groE* gene products in the synthesis of cellular DNA and RNA. *J. Bacteriol.* 157:694–96
229. Wada, M., Sekine, K., Itikawa, H. 1986. Participation of the *dnaK* and *dnaJ* gene products in phosphorylation of glutaminyl-tRNA synthetase and threonyl-tRNA synthetase of *Escherichia coli* K-12. *J. Bacteriol.* 168:213–20
230. Wadsworth, S. C. 1982. A family of related proteins is encoded by the major Drosophila heat shock gene family. *Mol. Cell. Biol.* 2:286–92
231. Walker, A. I., Hunt, T., Jackson, R. J., Anderson, C. W. 1985. Double-stranded DNA induces the phosphorylation of several proteins including the 90000 mol wt heat-shock protein in animal cell extracts. *EMBO J.* 4:139–45
232. Walter, G., Carbone, A., Welch, W. J. 1987. Medium tumor antigen of polymavirus transformation-defective mutant NG59 is associated with 73-kilodalton heat shock protein. *J. Virol.* 61:405–10
233. Wang, C., Lazarides, E. 1984. Arsenite-induced changes in methylation of the 70,000 dalton heat shock proteins in chicken embryo fibroblasts. *Biochem. Biophys. Res. Commun.* 119:735–43
234. Watowich, S. S., Morimoto, R. J. 1988. Complex regulation of heat shock- and glucose-responsive genes in human cells. *Mol. Cell. Biol.* 8:393–405
235. Watson, K., Dunlop, G., Cavicchioli, R. 1984. Mitochondrial and cytoplasmic protein syntheses not required for heat shock acquisition of ethanol- and thermotolerance in yeast. *FEBS Lett.* 169: 267–73
236. Welch, W. J., Feramisco, J. R. 1984. Nuclear and nucleolar localization of the 72,000-dalton heat shock in heat-shocked mammalian cells. *J. Biol. Chem.* 259: 4501–13
237. Welch, W. J., Feramisco, J. R. 1985. Rapid purification of mammalian 70,000-dalton stress proteins: affinity of the proteins for nucleotides. *Mol. Cell. Biol.* 5:1229–37
238. Welch, W. J., Suhan, J. P. 1985. Morphological study of the mammalian stress response: characterization of changes occurring in intracellular membraneous organelles, cytoskeletal elements, nucleus and nucleoli and appearance of intranuclear actin containing inclusion bodies in rat fibroblasts following heat shock. *J. Cell Biol.* 101:1198–1211
239. Welch, W. J., Suhan, J. P. 1986. Cellular and biochemical events in mammalian cells during and after recovery from physiological stress. *J. Cell Biol.* 103:2035–52
240. Werner-Washburne, M., Stone, D. E., Craig, E. A. 1987. Complex interactions among members of an essential subfamily of hsp70 genes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 7:2568–77
241. White, F., Currie, R. W. 1982. A mammalian response to trauma: the synthesis of a 71-kd protein. See Ref. 123, pp 379–86
242. Wu, C. 1980. The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286:854–60
243. Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., et al. 1987. Purification and properties of *Drosophila* heat shock activator protein. *Science* 238:1247–53
244. Yahara, I., Iida, H., Koyasu, S. 1986. A heat shock-resistant variant of Chinese hamster cell line constitutively expressing heat shock protein of Mr 90,000 at high level. *Cell Struct. Funct.* 11:65–73
245. Yamamori, T., Yura, T. 1982. Genetic control of heat-shock protein synthesis and its bearing on growth and thermal resistence in *Escherichia coli* K-12. *Proc. Natl. Acad. Sci. USA* 79:860–64
246. Yamamoto, T., McIntyre, J., Sell, S. M., Georgopoulos, C., Skowyra, D., Zylicz, M. 1987. Enzymology of the pre-priming steps in lambda DNA replication *in vitro*. *J. Biol. Chem.* 262:7996–99
247. Young, D., Lathigra, R., Mehlert, A. 1988. Stress-induced proteins as antigens in infectious diseases. *UCLA Symposium on Molecular & Cellular Biology, Vol. 96 Stress-Induced Proteins*, ed. J. R. Seramisco, S. L. Lindquist, M. L. Pardue, New York: Liss, p. 294
248. Zakeri, Z. F., Wolgemath, D., Hunt, C. R. 1988. Identification and analysis of a new member of the mouse Hsp70 gene family and characterization of its unique cellular and developmental pattern of expression in the male germ line. *Mol. Cell. Biol.* 8:2925–32
249. Zakeri, Z., Wolgemath, D. 1987. Developmental-stage specific expression of the hsp70 gene family during differentiation of the mammalian male germ line. *Mol. Cell. Biol.* 7:1791–96
250. Ziemiecki, A. 1986. Characterization of the monomeric and complex-associated forms of the *gag-onc* fusion proteins of three isolates of feline sarcoma virus: phosphorylation, kinase activity, acylation, and kinetics of complex formation. *Virology* 151:265–73
251. Ziemiecki, A., Catelli, M. G., Joab, I., Moncharmont, B. 1986. Association of the heat shock protein hsp90 with steroid hormone receptors and tyrosine kinase oncogene products. *Biochem. Biophys. Res. Commun.* 138:1298–1307
252. Zimmerman, J. L., Petri, W. L., Meselson, M. 1983. Accumulation of specific subsets of *D. melanogaster* heat shock mRNAs in normal development without heat shock. *Cell* 32:1161–70
253. Zylicz, M., Ang, D., Georgopoulos, C. 1987. The *grpE* protein of *Escherichia coli*. *J. Biol. Chem.* 262:17437–42
254. Zylicz, M., Georgopoulos, C. P. 1984. Purification and properties of the *Escherichia coli dnaK* replication protein. *J. Biol. Chem.* 259:8820–25
255. Zylicz, M., LeBowitz, J. H., McMacken, R., Georgopoulos, C. P. 1983. The *dnaK* protein of *Escherichia coli* possesses an ATPase and autophosphorylating activity and is essential to an *in vitro* DNA replication system. *Proc. Natl. Acad. Sci. USA* 80:6431–35
256. Picard, D., Salser, S. J., Yamamoto, K. R. 1988. A moveable and regulable inactivation function within the steroid-binding domain of the glucocorticoid receptor. *Cell*. In press
257 Bonner, J. J., Parks, C., Parker-Thornberg, J., Mortin, M. A., Pelham, H. R. B. 1984. The use of promoter fusions in Drosophila genetics. *Cell* 37:979–91
258 Kase, K. R., Hahn, G. M. 1975. Differential heat responses of normal and transformed human cells in tissue culture. *Nature* 255:288–90
259 Nelson, D. R., Kailleen, K. P. 1986. Heat shock proteins of vegetative and fruiting cells. *J. Bact.* 168:1100–06

EXHIBIT D

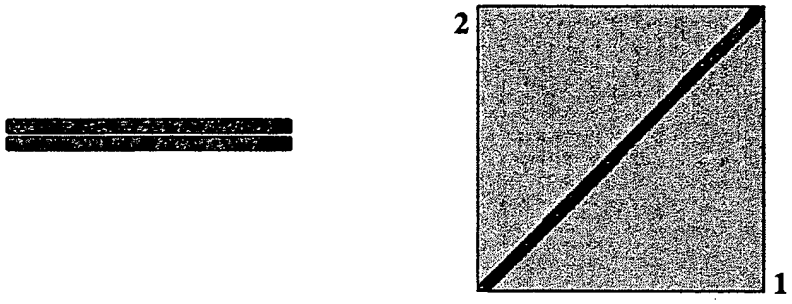
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 34862435	similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) [Rattus norvegicus]	**Length** 804 (1 .. 804)
Sequence 2	gi 15030324	Tumor rejection antigen gp96 [Mus musculus]	**Length** 802 (1 .. 802)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1550 bits (4014), Expect = 0.0
Identities = 788/805 (97%), Positives = 796/805 (97%), Gaps = 4/805 (0%)
```

```
Query: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALAGNEELTVKIKCDREKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENALAGNEELTVKIKCD+EKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALAGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDETAQ 300
           LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDE A+
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDEAAK 300

Query: 301 EEKEEADDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEE+DDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
```

```
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKSKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEK+KESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRVKEDEDDKTVMDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRR+KEDEDDKTVMDLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVMDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTTEDTTDDSEQDE-E 779
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDT+E+  +DSEQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTSEE-AEDSEQDEGE 779

Query: 780 ETDAGAEEEEEEQETEKEPTEKDEL 804
           E DAG EEEEE  ETEKE TEKDEL
Sbjct: 780 EMDAGTEEEEE--ETEKESTEKDEL 802

CPU time:     0.05 user secs.       0.00 sys. secs          0.05 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7160
Number of extensions: 3828
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 804
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 666
Effective length of database: 666,719,727
```

```
Effective search space: 444035338182
Effective search space used: 444035338182
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

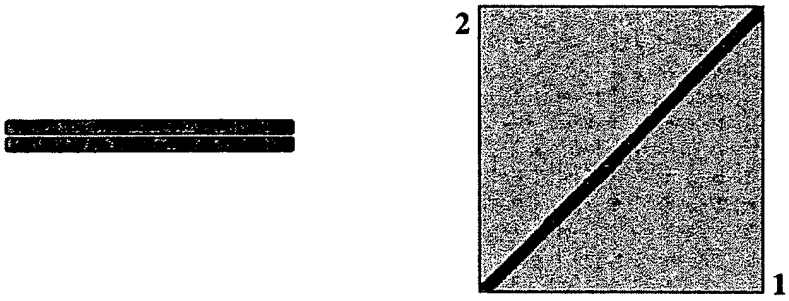
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1 gi 44890631 Tumor rejection antigen (gp96) 1 [Homo sapiens] **Length** 803 (1 .. 803)
Sequence 2 gi 15030324 Tumor rejection antigen gp96 [Mus musculus] **Length** 802 (1 .. 802)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1538 bits (3983), Expect = 0.0
Identities = 775/803 (96%), Positives = 791/803 (97%), Gaps = 1/803 (0%)
```

```
Query: 1   MRALWVLGLCCVLLTFGSVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MR LWVLGLCCVLLTFG VRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALSGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENAL+GNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALAGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTETVEEPMEEEEAAK 300
           LGRGTTITLVLKEEASDYLELDTIKNLV+KYSQFINFPIYVWSSKTETVEEP+EE+EAAK
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDEAAK 300

Query: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
```

```
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADDKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIAD+KY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHH TDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEKTKESREA EKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTV+DLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVMDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDEDE 780
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDP+A+VEEEPEEEPE+T+E+  ED+EQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTSEE-AEDSEQDEGE 779

Query: 781 EMDVGTDEEEETAKESTAEKDEL 803
           EMD GT+EEEE  ++ + EKDEL
Sbjct: 780 EMDAGTEEEEEETEKESTEKDEL 802

CPU time:     0.05 user secs.       0.01 sys. secs          0.06 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7092
Number of extensions: 3817
Number of successful extensions: 61
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
```

```
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

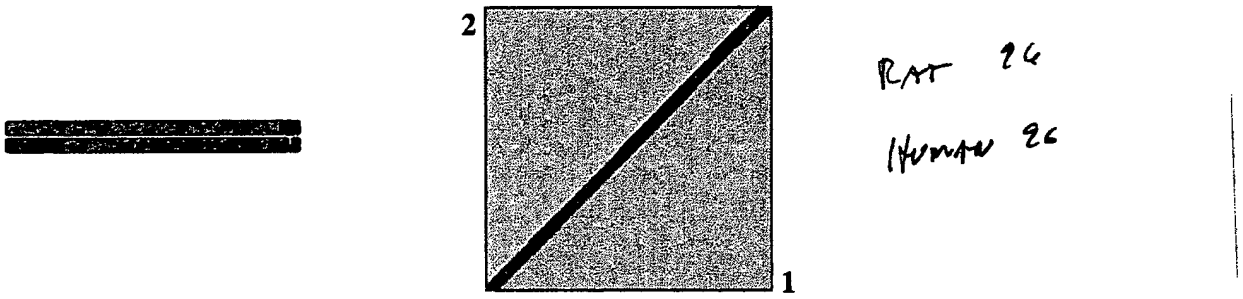
NCBI Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 34862435	similar to Endoplasmin precursor (Endoplasmic reticulum protein 99) (94 kDa glucose-regulated protein) (GRP94) (ERP99) (Polymorphic tumor rejection antigen 1) (Tumor rejection antigen gp96) [Rattus norvegicus]	**Length** 804 (1 .. 804)
Sequence 2	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1527 bits (3954), Expect = 0.0
Identities = 773/806 (95%), Positives = 790/806 (97%), Gaps = 5/806 (0%)
```

```
Query: 1   MRVLWVLGLCCVLLTFGFVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60
           MR LWVLGLCCVLLTFG VRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG
Sbjct: 1   MRALWVLGLCCVLLTFGSVRADDEVDVDGTVEEDLGKSREGSRTDDEVVQREEEAIQLDG 60

Query: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120
           LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL
Sbjct: 61  LNASQIRELREKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISL 120

Query: 121 TDENALAGNEELTVKIKCDREKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180
           TDENAL+GNEELTVKIKCD+EKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE
Sbjct: 121 TDENALSGNEELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTE 180

Query: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240
           AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT
Sbjct: 181 AQEDGQSTSELIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNT 240

Query: 241 LGRGTTITLVLKEEASDYLELDTIKNLVRKYSQFINFPIYVWSSKTETVEEPLEEDETAQ 300
           LGRGTTITLVLKEEASDYLELDTIKNLV+KYSQFINFPIYVWSSKTETVEEP+EE+E A+
Sbjct: 241 LGRGTTITLVLKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKTETVEEPMEEEEAAK 300

Query: 301 EEKEEADDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360
           EEKEE+DDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK
```

```
Sbjct: 301 EEKEESDDEAAVEEEEEEKKPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEYKAFYK 360

Query: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420
           SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD
Sbjct: 361 SFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRRVFITD 420

Query: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADEKY 480
           DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIAD+KY
Sbjct: 421 DFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKIADDKY 480

Query: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHSTDITSLDQYVERMKEKQDKIYF 540
           NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHH TDITSLDQYVERMKEKQDKIYF
Sbjct: 481 NDTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEKQDKIYF 540

Query: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600
           MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE
Sbjct: 541 MAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKEGVKFDE 600

Query: 601 SEKSKESREATEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660
           SEK+KESREA EKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER
Sbjct: 601 SEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGWSGNMER 660

Query: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRVKEDEDDKTVMDLAVVLFET 720
           IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRR+KEDEDDKTV+DLAVVLFET
Sbjct: 661 IMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLAVVLFET 720

Query: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPEAQVEEEPEEEPEDTTEDTTDDSEQDE-E 779
           ATLRSGYLLPDTKAYGDRIERMLRLSLNIDP+A+VEEEPEEEPE+T EDTT+D+EQDE E
Sbjct: 721 ATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDEDE 780

Query: 780 ETDAGAEEEEEEQETEKEPT-EKDEL 804
           E D G +EEE   ET KE T EKDEL
Sbjct: 781 EMDVGTDEEE---ETAKESTAEKDEL 803

CPU time:     0.05 user secs.       0.01 sys. secs          0.06 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 7146
Number of extensions: 3810
Number of successful extensions: 63
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 804
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 666
Effective length of database: 666,719,727
```

```
Effective search space: 444035338182
Effective search space used: 444035338182
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

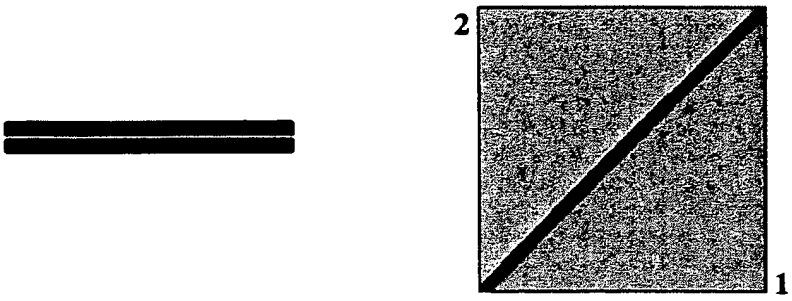
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

Score = 1436 bits (3717), Expect = 0.0
Identities = 725/733 (98%), Positives = 731/733 (98%), Gaps = 1/733 (0%)

```
Query:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
              MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
Conflict  7         *
modified  7         *
modified  5       *
HSPCA     1   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   61   YETLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
              YE+LTDPSKLDSGKELHINLIP+KQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct:   61   YESLTDPSKLDSGKELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
HSPCA    61   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   121  ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
              ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct:   121  ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
HSPCA    121  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   181  GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKED 240
              GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKE+
Sbjct:   181  GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
modified 231                                                        *
HSPCA    181  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
```

```
Query:    241 KEEEKEKEEKESEDKPEIEDVGSDEEEE-KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 299
              KEEEKEKEEKES+DKPEIEDVGSDEEEE KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
modified  263                          *
Conflict  243   ****
HSPCA     241 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    300 NPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 359
              NPDDITNEEYGEFYKSLTNDWE+HLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
Conflict  356                                                        *
HSPCA     301 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    360 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 419
              NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
HSPCA     361 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    420 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 479
              CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
HSPCA     421 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    480 YCTRMKENQKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 539
              YCTRMKENQKHIY+ITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
HSPCA     481 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    540 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 599
              TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
HSPCA     541 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    600 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 659
              VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
HSPCA     601 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    660 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPL 719
              KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPT DDTSAAVTEEMPPL
Sbjct:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
HSPCA     661 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    720 EGDDDTSRMEEVD 732
              EGDDDTSRMEEVD
Sbjct:    721 EGDDDTSRMEEVD 733
HSPCA     721 +++++++++++++

CPU time:      0.04 user secs.       0.00 sys. secs          0.04 total secs.

Lambda      K       H
   0.312    0.131    0.361

Gapped
Lambda      K       H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

```
Number of Sequences: 1
Number of Hits to DB: 5682
Number of extensions: 3005
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

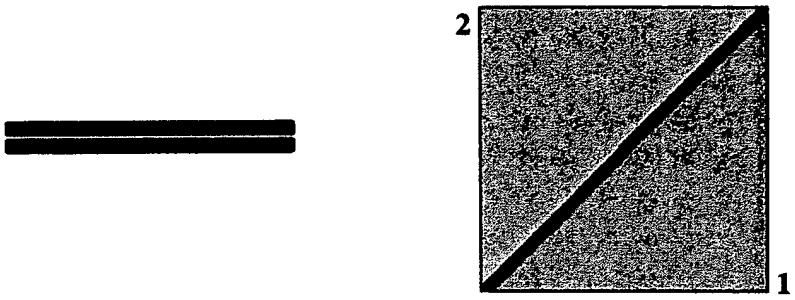
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 48734827	Heat shock protein 1, alpha [Rattus norvegicus]	**Length** 733 (1 .. 733)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1450 bits (3754), Expect = 0.0
Identities = 732/733 (99%), Positives = 733/733 (99%)
```

```
Query:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
              MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct:    1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
Conflict  7         *
modified  7         *
modified  5       *
HSPCA     1   ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    61  YESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
              YESLTDPSKLDSGKELHINLIP+KQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct:    61  YESLTDPSKLDSGKELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
HSPCA     61  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
              ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct:    121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
HSPCA     121 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
              GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE
Sbjct:    181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240
modified  231                                          *
HSPCA     181 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
```

```
Query:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
              KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct:    241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300
modified 263                          *
Conflict 243   ****
HSPCA     241 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
              NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct:    301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
Conflict 356                                                          *
HSPCA     301 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
              NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct:    361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420
HSPCA     361 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
              CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct:    421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480
HSPCA     421 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
              YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct:    481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540
HSPCA     481 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
              TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct:    541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600
HSPCA     541 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
              VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct:    601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660
HSPCA     601 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
              KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL
Sbjct:    661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720
HSPCA     661 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    721 EGDDDTSRMEEVD 733
              EGDDDTSRMEEVD
Sbjct:    721 EGDDDTSRMEEVD 733
HSPCA     721 +++++++++++++

CPU time:      0.04 user secs.       0.01 sys. secs          0.05 total secs.

Lambda     K      H
   0.312    0.131    0.360

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
```

```
Number of Sequences: 1
Number of Hits to DB: 5739
Number of extensions: 3027
Number of successful extensions: 71
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 733
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 596
Effective length of database: 666,719,728
Effective search space: 397364957888
Effective search space used: 397364957888
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

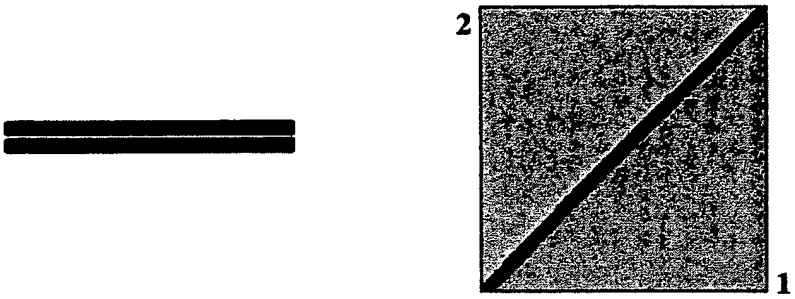
Blast 2 Sequences results

PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 48734827	Heat shock protein 1, alpha [Rattus norvegicus]	**Length** 733 (1 .. 733)



NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score = 1438 bits (3722), Expect = 0.0
Identities = 726/733 (99%), Positives = 731/733 (99%), Gaps = 1/733 (0%)
```

```
Query: 1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60
           MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR
Sbjct: 1   MPEETQTQDQPMEEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIR 60

Query: 61  YETLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120
           YE+LTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME
Sbjct: 61  YESLTDPSKLDSGKELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFME 120

Query: 121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180
           ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM
Sbjct: 121 ALQAGADISMIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPM 180

Query: 181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKED 240
           GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKE+
Sbjct: 181 GRGTKVILHLKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEE 240

Query: 241 KEEEKEKEEKESEDKPEIEDVGSDEEEE-KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 299
           KEEEKEKEEKES+DKPEIEDVGSDEEEE KKDGDKKKKKKIKEKYIDQEELNKTKPIWTR
Sbjct: 241 KEEEKEKEEKESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTR 300

Query: 300 NPDDITNEEYGEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 359
           NPDDITNEEYGEFYKSLTNDWE+HLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN
Sbjct: 301 NPDDITNEEYGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFENRKKKN 360
```

```
Query: 360 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 419
           NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK
Sbjct: 361 NIKLYVRRVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKK 420

Query: 420 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 479
           CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD
Sbjct: 421 CLELFTELAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKD 480

Query: 480 YCTRMKENQKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 539
           YCTRMKENQKHIY+ITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK
Sbjct: 481 YCTRMKENQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGK 540

Query: 540 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 599
           TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI
Sbjct: 541 TLVSVTKEGLELPEDEEEKKKQEEKKTKFENLCKIMKDILEKKVEKVVVSNRLVTSPCCI 600

Query: 600 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 659
           VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV
Sbjct: 601 VTSTYGWTANMERIMKAQALRDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSV 660

Query: 660 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPL 719
           KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPT DDTSAAVTEEMPPL
Sbjct: 661 KDLVILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPL 720

Query: 720 EGDDDTSRMEEVD 732
           EGDDDTSRMEEVD
Sbjct: 721 EGDDDTSRMEEVD 733

CPU time:     0.06 user secs.       0.00 sys. secs          0.06 total secs.

Lambda      K      H
   0.312    0.131    0.361

Gapped
Lambda      K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5690
Number of extensions: 3010
Number of successful extensions: 65
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 1
Number of HSP's successfully gapped: 1
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
```

X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)

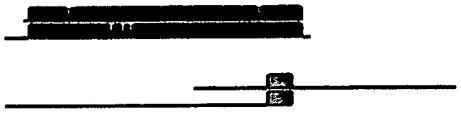
Blast 2 Sequences results

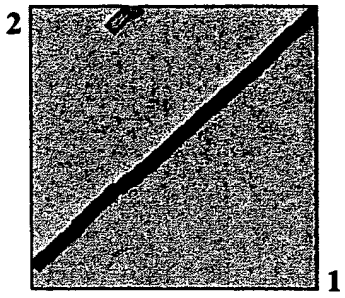
PubMed Entrez BLAST OMIM Taxonomy Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 40254816	heat shock 90kDa protein 1, alpha; heat shock 90kD protein 1, alpha [Homo sapiens]	**Length** 732 (1 .. 732)
Sequence 2	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)



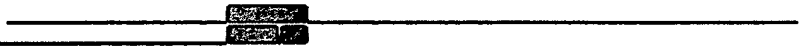


NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score =  638 bits (1645), Expect = 0.0
Identities = 344/729 (47%), Positives = 479/729 (65%), Gaps = 27/729 (3%)
```

```
Query: 15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYETLTDPSKLDSGK 74
           E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  +LTD + L   +
Sbjct: 71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130

Query: 75  ELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129
           EL + +  +K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct: 131 ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190

Query: 130 MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189
           +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct: 191 LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250

Query: 190 LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEEKEKEE 249
           LKE+ ++YLE   IK +VKK+SQFI +PI ++  K          E  E+  +EEE  KEE
Sbjct: 251 LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT--------ETVEEPMEEEEAAKEE 302

Query: 250 KESEDKPEIEDVGSDEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEEY 309
           KE  D    ++   +EEEE+K   K K KK+++   D E +N  KPIW R   ++  +EY
Sbjct: 303 KEESD----DEAAVEEEEEEK---KPKTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDEY 355

Query: 310 GEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVRR 367
             FYKS + + +D +A  HF+ EG++ F++LFVP  AP  LF+    KK + IKLYVRR
Sbjct: 356 KAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVRR 415
```

```
Query: 368 VFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTEL 427
           VFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  ++
Sbjct: 416 VFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKKI 475

Query: 428 AEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKEN 487
           A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct: 476 ADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKEK 534

Query: 488 QKHIYYITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTKE 547
           Q  IY++ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V KE
Sbjct: 535 QDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAKE 594

Query: 548 GLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYGW 606
           G++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YGW
Sbjct: 595 GVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYGW 654

Query: 607 TANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDLV 663
           + NMERIMKAQA    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct: 655 SGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDLA 714

Query: 664 ILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTADDTSAAVTEEMPPLEGDD 723
           ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++       E       D
Sbjct: 715 VLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDT 774

Query: 724 DTSRMEEVD 732
           +    EE+D
Sbjct: 775 EQDEDEEMD 783

Score = 35.4 bits (80), Expect = 8.6
Identities = 21/73 (28%), Positives = 38/73 (51%), Gaps = 2/73 (2%)
```



```
Query: 208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEDKEEEKEKEEKESEDKPEIEDVGSDEEE 267
           K +   I   + L +  + D +V ++  EE E+  E+  ++ ++ ED  E  DVG+DEEE
Sbjct: 733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790

Query: 268 EKKDGDKKKKKKI 280
           E       +K ++
Sbjct: 791 ETAKESTAEKDEL 803

CPU time:     0.08 user secs.       0.00 sys. secs          0.08 total secs.

Lambda     K      H
   0.312    0.131    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5882
Number of extensions: 3409
Number of successful extensions: 66
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 5
```

```
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 732
Length of database: 666,719,865
Length adjustment: 137
Effective length of query: 595
Effective length of database: 666,719,728
Effective search space: 396698238160
Effective search space used: 396698238160
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

Blast 2 Sequences results

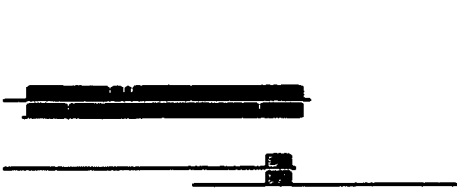
PubMed Entrez BLAST OMIM Taxonomy Structure

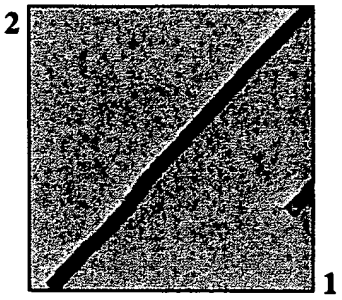
BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1 gi 44890631 Tumor rejection antigen (gp96) 1 [Homo sapiens] **Length** 803 (1 .. 803)

Sequence 2 gi 48734827 Heat shock protein 1, alpha [Rattus norvegicus] **Length** 733 (1 .. 733)





NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score =  641 bits (1654), Expect = 0.0
Identities = 346/730 (47%), Positives = 482/730 (65%), Gaps = 28/730 (3%)
```

```
Query: 71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130
           E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  SLTD + L   +
Sbjct: 15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGK 74

Query: 131 ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190
           EL + +  +K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct: 75  ELHINLIPNKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129

Query: 191 LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250
           +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct: 130 MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189

Query: 251 LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT-------ETVEEPMEEEEAAKEEK 303
           LKE+ ++YLE   IK +VKK+SQFI +PI ++  K         +  EE  E+EE  ++E+
Sbjct: 190 LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEE 249

Query: 304 EESDDEAAVE-----EEEEEKKP----KTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDE 354
           +ESDD+  +E     EEEEEKK      K KK+++   D E +N  KPIW R   ++  +E
Sbjct: 250 KESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEE 309

Query: 355 YKAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVR 414
           Y  FYKS + + ++ +A  HF+ EG++ F+++LFVP  AP  LF+    KK + IKLYVR
Sbjct: 310 YGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVR 367

Query: 415 RVFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKK 474
           RVFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  +
Sbjct: 367 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFSE 426
```

```
Sbjct: 368 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTE 427

Query: 475 IADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKE 533
           +A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct: 428 LAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKE 487

Query: 534 KQDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAK 593
            Q  IYF+ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V K
Sbjct: 488 NQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTK 547

Query: 594 EGVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYG 653
           EG++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YG
Sbjct: 548 EGLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYG 606

Query: 654 WSGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDL 713
           W+ NMERIMKAQA    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct: 607 WTANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDL 663

Query: 714 AVVLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTED 773
            ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++      E       D
Sbjct: 664 VILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPLEGD 723

Query: 774 TEQDEDEEMD 783
            +    EE+D
Sbjct: 724 DDTSRMEEVD 733

Score = 38.5 bits (88), Expect = 1.1
Identities = 22/70 (31%), Positives = 38/70 (53%), Gaps = 2/70 (2%)
```

```
Query: 733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790
           K +   I   + L +  + D +V ++  EE EE  E+  ++ ++ +D  E  DVG+DEEE
Sbjct: 208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEEKESDDKPEIEDVGSDEEE 267

Query: 791 ETAKESTAEK 800
           E  K+   +K
Sbjct: 268 EEKKDGDKKK 277

CPU time:     0.11 user secs.       0.01 sys. secs          0.12 total secs.

Lambda     K      H
   0.312    0.130    0.361

Gapped
Lambda     K      H
   0.267   0.0410    0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5960
Number of extensions: 3415
Number of successful extensions: 68
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 16
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
```

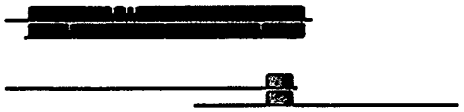
```
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

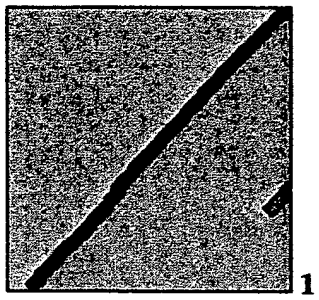
Blast 2 Sequences results

BLAST 2 SEQUENCES RESULTS VERSION BLASTP 2.2.9 [May-01-2004]

Matrix BLOSUM62 gap open: 11 gap extension: 1
x_dropoff: 50 expect: 10.000 wordsize: 3 Filter Align

Sequence 1	gi 44890631	Tumor rejection antigen (gp96) 1 [Homo sapiens]	**Length** 803 (1 .. 803)
Sequence 2	gi 1170384	Heat shock protein HSP 90-alpha (HSP 86) (Tumor specific transplantation 86 kDa antigen) (TSTA).	**Length** 733 (1 .. 733)



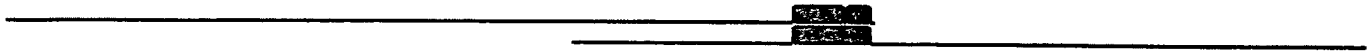


NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

```
Score =  641 bits (1653), Expect = 0.0
Identities = 346/730 (47%), Positives = 481/730 (65%), Gaps = 28/730 (3%)
```

```
Query:    71  EKSEKFAFQAEVNRMMKLIINSLYKNKEIFLRELISNASDALDKIRLISLTDENALSGNE 130
              E+ E FAFQAE+ ++M LIIN+ Y NKEIFLRELISN+SDALDKIR  SLTD + L   +
Sbjct:    15  EEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNSSDALDKIRYESLTDPSKLDSGK 74
HSPCA     15  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   131  ELTVKIKCDKEKNLLHVTDTGVGMTREELVKNLGTIAKSGTSEFLNKMTEAQEDGQSTSE 190
              EL + +   K+   L + DTG+GMT+ +L+ NLGTIAKSGT  F+    EA + G   S
Sbjct:    75  ELHINLIPSKQDRTLTIVDTGIGMTKADLINNLGTIAKSGTKAFM----EALQAGADIS- 129
HSPCA     75  ++++++++++++++++++++++++++++++++++++++++++++    ++++++++++

Query:   191  LIGQFGVGFYSAFLVADKVIVTSKHNNDTQHIWESDSNEFSVIADPRGNTLGRGTTITLV 250
              +IGQFGVGFYSA+LVA+KV V +KHN+D Q+ WES +     +    G  +GRGT + L
Sbjct:   130  MIGQFGVGFYSAYLVAEKVTVITKHNDDEQYAWESSAGGSFTVRTDTGEPMGRGTKVILH 189
HSPCA    130  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:   251  LKEEASDYLELDTIKNLVKKYSQFINFPIYVWSSKT-------ETVEEPMEEEEAAKEEK 303
              LKE+ ++YLE   IK +VKK+SQFI +PI ++  K        +  EE  E+EE  ++E+
Sbjct:   190  LKEDQTEYLEERRIKEIVKKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEE 249
HSPCA    190  ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified 231                                        *
Conflict 243                                                  ****

Query:   304  EESDDEAAVE-----EEEEEKKP----KTKKVEKTVWDWELMNDIKPIWQRPSKEVEEDE 354
              +ESDD+  +E     EEEEEKK     K KK+++   D E +N  KPIW R   ++  +E
```

```
Sbjct:    250 KESDDKPEIEDVGSDEEEEEKKDGDKKKKKKIKEKYIDQEELNKTKPIWTRNPDDITNEE 309
HSPCA     250 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified  263              *

Query:    355 YKAFYKSFSKESDDPMAYIHFTAEGEVTFKSILFVPTSAPRGLFDEYGSKKSDYIKLYVR 414
              Y  FYKS + + ++ +A  HF+ EG++ F+++LFVP  AP  LF+    KK + IKLYVR
Sbjct:    310 YGEFYKSLTNDWEEHLAVKHFSVEGQLEFRALLFVPRRAPFDLFEN--RKKKNNIKLYVR 367
HSPCA     310 ++++++++++++++++++++++++++++++++++++++++++++++  ++++++++++++
Conflict  356                                                 *

Query:    415 RVFITDDFHDMMPKYLNFVKGVVDSDDLPLNVSRETLQQHKLLKVIRKKLVRKTLDMIKK 474
              RVFI D+  +++P+YLNF++GVVDS+DLPLN+SRE LQQ K+LKVIRK LV+K L++  +
Sbjct:    368 RVFIMDNCEELIPEYLNFIRGVVDSEDLPLNISREMLQQSKILKVIRKNLVKKCLELFTE 427
HSPCA     368 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    475 IADDKYN-DTFWKEFGTNIKLGVIEDHSNRTRLAKLLRFQSSHHPTDITSLDQYVERMKE 533
              +A+DK N   F+++F  NIKLG+ ED  NR +L++LLR+ +S    ++ SL  Y  RMKE
Sbjct:    428 LAEDKENYKKFYEQFSKNIKLGIHEDSQNRKKLSELLRYYTSASGDEMVSLKDYCTRMKE 487
HSPCA     428 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    534 KQDKIYFMAGSSRKEAESSPFVERLLKKGYEVIYLTEPVDEYCIQALPEFDGKRFQNVAK 593
               Q  IYF+ G ++ +  +S FVERL K G EVIY+ EP+DEYC+Q L EF+GK   +V K
Sbjct:    488 NQKHIYFITGETKDQVANSAFVERLRKHGLEVIYMIEPIDEYCVQQLKEFEGKTLVSVTK 547
HSPCA     488 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    594 EGVKFDESEKTKESREAVEKEFEPLLNWMKDKALKDKIEKAVVSQRLTESPCALVASQYG 653
              EG++  E E+ K+ +E  + +FE L   MKD  L+ K+EK VVS RL  SPC +V S YG
Sbjct:    548 EGLELPEDEEEKKKQEEKKTKFENLCKIMKD-ILEKKVEKVVVSNRLVTSPCCIVTSTYG 606
HSPCA     548 ++++++++++++++++++++++++++++++ ++++++++++++++++++++++++++++++

Query:    654 WSGNMERIMKAQAYQTGKDISTNYYASQKKTFEINPRHPLIRDMLRRIKEDEDDKTVLDL 713
              W+ NMERIMKAQA    +D ST  Y + KK  EINP H +I  + ++ + D++DK+V DL
Sbjct:    607 WTANMERIMKAQAL---RDNSTMGYMAAKKHLEINPDHSIIETLRQKAEADKNDKSVKDL 663
HSPCA     607 ++++++++++++   ++++++++++++++++++++++++++++++++++++++++++++++

Query:    714 AVVLFETATLRSGYLLPDTKAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTED 773
               ++L+ETA L SG+ L D + + +RI RM++L L ID D    ++      E       D
Sbjct:    664 VILLYETALLSSGFSLEDPQTHANRIYRMIKLGLGIDEDDPTVDDTSAAVTEEMPPLEGD 723
HSPCA     664 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

Query:    774 TEQDEDEEMD 783
               +    EE+D
Sbjct:    724 DDTSRMEEVD 733
HSPCA     724 ++++++++++

Score = 38.5 bits (88), Expect = 1.1
Identities = 22/70 (31%), Positives = 38/70 (53%), Gaps = 2/70 (2%)
```



```
Query:    733 KAYGDRIERMLRLSLNIDPDAKVEEEPEEEPEETAEDTTEDTEQDED--EEMDVGTDEEE 790
              K +   I   + L +  + D +V ++  EE EE  E+  ++ ++ +D  E  DVG+DEEE
Sbjct:    208 KKHSQFIGYPITLFVEKERDKEVSDDEAEEKEEKEEEKEKEEKESDDKPEIEDVGSDEEE 267
Conflict  243                                    ****
modified  231                        *
HSPCA     208 ++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++
modified  263                                                      *

Query:    791 ETAKESTAEK 800
              E  K+   +K
```

```
Sbjct:    268 EEKKDGDKKK 277
HSPCA     268 ++++++++++

CPU time:      0.10 user secs.        0.00 sys. secs          0.10 total secs.

Lambda      K       H
   0.312     0.130     0.361

Gapped
Lambda      K       H
   0.267    0.0410     0.140


Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Sequences: 1
Number of Hits to DB: 5958
Number of extensions: 3413
Number of successful extensions: 68
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 1
Number of HSP's gapped: 16
Number of HSP's successfully gapped: 2
Number of extra gapped extensions for HSPs above 10.0: 0
Length of query: 803
Length of database: 666,719,865
Length adjustment: 138
Effective length of query: 665
Effective length of database: 666,719,727
Effective search space: 443368618455
Effective search space used: 443368618455
Neighboring words threshold: 9
Window for multiple hits: 0
X1: 16 ( 7.2 bits)
X2: 129 (50.0 bits)
X3: 129 (50.0 bits)
S1: 42 (22.0 bits)
S2: 80 (35.4 bits)
```

EXHIBIT H

Protein folding in the cell

Mary-Jane Gething & Joseph Sambrook

In the cell, as *in vitro*, the final conformation of a protein is determined by its amino-acid sequence. But whereas some isolated proteins can be denatured and refolded *in vitro* in the absence of other macromolecular cellular components, folding and assembly of polypeptides *in vivo* involves other proteins, many of which belong to families that have been highly conserved during evolution.

UNTIL recently, scientists using biophysical techniques to study refolding of polypeptides *in vitro* had little need for communication with those investigating the biosynthesis and maturation of proteins within cells. But the realization that the attainment of correct tertiary and quaternary structure is an important determinant of efficient intracellular protein transport[1-3] led to the development of techniques to analyse the early stages of protein folding *in vivo*[4]. These studies have shown that in cells families of abundant proteins modulate and promote protein folding, assembly and disassembly, and facilitate the degradation of malfolded polypeptides. Here we review current knowledge of these proteins and discuss current theories of the mechanisms by which they function.

Protein folding *in vitro* and *in vivo*

Although the three-dimensional structures of several hundred proteins are now known in great detail[5,6], the pathways by which such polypeptides attain their native configurations remain substantially undefined. Anfinsen's classic experiments on the refolding of ribonuclease *in vitro*[7,8] established that all the information required to determine the final conformation of a protein can reside in the polypeptide chain itself: the denatured enzyme can refold into its native conformation in the absence of other proteins. Similar results have since been obtained with several other small, single-domain polypeptides, and with a few larger, more complex proteins (reviewed in refs 9-13). Such studies suggest that refolding *in vitro* may be initiated by (1) collapse of hydrophobic regions into the interior of the molecule, (2) formation of stable secondary structures that provide a framework for subsequent folding, and (3) formation of covalent interactions, such as disulphide bonds, that stabilize the polypeptide in particular conformations. Evidence has been obtained in support of each of these mechanisms and it is likely that all three may operate in conjunction during the early stages of refolding. Subsequent folding seems to occur through a limited number of pathways involving distinct intermediates ('molten globules'[14,15] or 'compact intermediates'[13]) that have significant secondary structure and a compact form but lack a well-defined tertiary structure and expose more hydrophobic surface than fully folded molecules. These intermediates seem to be in rapid equilibrium with the fully denatured state and are only slowly converted to the native state. Thus the rate-limiting step in the refolding process frequently occurs at a very late stage, just before the protein adopts its final, native conformation.

Despite their value in defining the types of intramolecular interactions that drive polypeptide folding, *in vitro* experiments do not accurately reflect the process of folding of nascent proteins in the interior of a cell. Refolding *in vitro* is frequently very inefficient in comparison to folding *in vivo*, and often requires protein concentrations and physicochemical conditions very different from these occurring intracellularly. Furthermore, although refolding experiments involve the whole polypeptide chain, the opportunity exists *in vivo* for folding to commence as soon as the N-terminal portion of the nascent chain emerges from the ribosome (or from the lipid bilayer following membrane translocation). Finally, many proteins exist in cells as homo- or hetero-oligomeric complexes that in some cases are assembled before folding of the individual polypeptide chains is complete. The probability is slight that such complexes could form *in vitro* at the low subunit concentrations occurring intracellularly[16]; the probability is even lower that they could form before individual chains have become irreversibly malfolded.

To investigate protein folding *in vivo* it was first necessary to devise assays for polypeptide conformation that do not depend on obtaining large quantities of partially folded proteins that are sufficiently pure for physicochemical measurements. The first such assays analysed the formation of disulphide bonds during the folding of biosynthetically labelled secretory proteins and showed that these bonds can form even before synthesis of a polypeptide is completed, and that disulphide bond formation occurs *in vivo* at a significantly faster rate than can be achieved under the most favourable conditions *in vitro*[17]. Subsequently, additional *in vivo* folding assays were developed that use conformation-specific antibodies, protease sensitivity, or sucrose density gradient centrifugation to probe the tertiary and quaternary structure of radiolabelled proteins (reviewed in ref. 4). Studies using these techniques showed that although individual domains of a nascent polypeptide may fold very rapidly, acquisition of the final native structure of the whole protein can proceed comparatively slowly. Furthermore, partially folded intermediates whose structures would seem unlikely to be stable in the presumably aqueous environment of the cell's interior can lie dormant for many minutes to many hours before folding is completed[3,18,19]. Despite such extended pauses in the assembly process, folding *in vivo* of wild-type proteins is usually highly efficient with >95% of the newly synthesized polypeptides eventually attaining their native three-dimensional structures[3,20]. Polypeptide misfolding and aggregation, frequently a major problem during refolding *in vitro*[21], rarely occurs *in vivo* except with mutant proteins or during synthesis at elevated temperatures. Finally, partially folded polypeptides can frequently be isolated as complexes with specific cellular proteins, notably members of stress protein families[3,22-26].

The rest of this article summarizes the results of studies that have revealed the existence of at least two classes of proteins involved in polypeptide folding in cells. The first class includes conventional enzymes that catalyse specific isomerization steps that may otherwise limit the rate of folding of some proteins, whereas a second class of 'chaperones' stabilize unfolded or partially folded structures and prevent the formation of inappropriate intra- or interchain interactions. Some members of this second class also interact with apparently native protein molecules to promote rearrangement of protein–protein interactions in oligomeric structures.

Enzymes involved in protein folding

In vitro, two rate-determining steps involving isomerization of covalent bonds can be catalysed by purified cellular enzymes. Protein disulphide isomerase (PDI) catalyses thiol/disulphide interchange reactions and, depending on the nature of the polypeptide substrate and the imposed redox potential, promotes protein disulphide formation, isomerization or reduction[27,28]. PDI does not determine the polypeptide's folding pathway, but rather facilitates formation of the correct set of disulphide bonds

by promoting rapid reshuffling of incorrect disulphide pairings. Proteins with peptidyl prolyl *cis-trans* isomerase (PPIase) activity catalyse the otherwise slow isomerization of X-P peptide bonds (where X is any amino acid and P is proline in single-letter amino-acid code) and can accelerate the refolding of proline-containing polypeptides *in vitro*[10,29,30] and *in vivo*[31].

PDI and the thioredoxin-like proteins

Several lines of evidence suggest that PDI activity is required for folding of nascent polypeptides in the endoplasmic reticulum (ER) of eukaryotic cells. Thus the abundance of the enzyme in the ER of different cell types correlates with the level of secretory protein synthesis[17] and *in vivo* chemical crosslinking studies demonstrate a specific association between the enzyme and newly synthesized immunoglobulin chains in the ER (ref. 32). Finally, reintroduction of purified PDI into microsomes evacuated of their lumenal content by alkali or detergent treatment restores cotranslational formation of disulphide bonds in proteins synthesized using a cell-free system[33]. PDI is an essential protein in yeast[34,35], but the biological process whose disruption leads to lethality has not yet been defined.

Mammalian PDI is a dimer of identical subunits of relative molecular mass 57,000 (M_r 57K), each of which contains duplications of domains showing strong homology to thioredoxin[36], a small redox protein present in all classes of organisms from bacteria to higher eukaryotes[37]. Computer modelling studies based on the known three-dimensional structure of *Escherichia coli* thioredoxin indicate that a functional PDI dimer contains four thioredoxin-like domains each having a dithiol/disulphide active site located on a prominent loop at the surface of the molecule[38].

PDI's role in the ER may not be limited to disulphide isomerization[28]. First, many types of mammalian cells contain, in addition to substantial amounts of homodimeric PDI, the enzyme prolyl-4-hydroxylase which causes extensive modification of proline residues in nascent collagen molecules. This enzyme, an $\alpha_2\beta_2$ tetramer whose β-chain dimers are identical to PDI, also has PDI activity (reviewed in ref. 27). The 64K α-chains form the binding site for the peptide substrate to be hydroxylated, but it has not yet been determined whether the dithiol/disulphide active sites of the PDI/β-subunits are directly involved in the hydroxylation reaction. Second, the 'glycosylation site binding protein' (GSBP) component identified using a glycosylation site photoaffinity probe[39] is identical to PDI (ref. 34). However, involvement of this enzyme during *N*-linked glycosylation in the ER is not proved, as depletion of PDI from microsomes does not affect their capacity to support oligosaccharide addition to nascent polypeptides[40]. Finally, PDI has also been identified as a component of the microsomal triglyceride transfer protein complex[41].

Recent studies have identified other ER proteins containing thioredoxin homology units. ERp59, ERp61 and ERp72, three members of a set of proteins whose synthesis is induced at onset of immunoglobulin secretion in murine B cells[42], have been characterized by complementary DNA cloning and sequencing. ERp59 is identical to PDI (ref. 43). Although ERp61 contains two thioredoxin-like domains found in the same relative positions as in the PDI molecule, the rest of its sequence is unrelated to that of PDI (R. Mazzarella and M. Green, personal communication). ERp72 contains three thioredoxin homology units, two of which are spaced as in PDI, embedded in otherwise unrelated sequences[43]. Neither ERp61 nor ERp72 have PDI activity *in vitro*. Finally, an essential gene, *EUG1*, encoding another PDI-related ER protein containing thioredoxin homology units but lacking PDI activity, has been identified in *Saccharomyces cerevisiae* (C. Tachibana and T. Stevens, personal communication). Whether the unknown functions of these proteins use the redox activity of their thioredoxin-related structural units is still in question. In this context it is of interest that thioredoxin itself is required for assembly of filamentous phages in *E. coli*, playing a part that does not involve its redox activity as site-specific mutation of either or both of the active site cysteines does not alter the ability of the mutant proteins to support phage assembly (for review, see ref. 44). Thioredoxin is thought to confer processivity on the reaction that leads to the displacement of the intracellular phage protein pV from single-stranded phage DNA and to its replacement at the membrane by the major coat protein, pVIII.

Thus the ER houses an extended family of enzymes (Table 1) that may use thioredoxin-like domains containing dithiol/disulphide active sites to carry out various functions in the co- and post-translational modification of secretory proteins. Whether these proteins may also have roles in protein assembly that do not involve redox activity, and whether there may be, in addition to thioredoxin itself, other members of this family located in other compartments of the cell remains to be determined.

Peptidyl prolyl *cis–trans* isomerases

Proteins with PPIase activity are highly abundant and widely distributed, being found in virtually all tissues and organisms, from bacteria to mammals (reviewed in ref. 10). Those proteins characterized so far fall into two structurally unrelated families (Table 1), which are named after the clinically important immunosuppressive agents that inhibit their isomerase activity. Thus the eukaryotic cyclophilin proteins bind cyclosporin A (CsA) with high affinity, whereas the FK506-binding proteins bind the structurally distinct compounds FK506 and rapamycin (reviewed in ref. 45). Both CsA and FK506 mediate their immunosuppressive action by preventing the transcription of genes involved in activation of T lymphocytes, whereas rapamycin potently inhibits the response of T cells to the lymphokine IL-2 (ref. 45). These immunosuppressive drugs do not act through inhibition of the PPIase activity of T cells as they are effective at concentrations far below those of the PPIase enzymes, and they inhibit distinct signalling pathways. Rather it seems that cyclophilin and FKBP bind the drugs, which are cyclic peptides, and present them (to as yet undefined targets) in a bioactive conformation that may require a *cis–trans* isomerization around one of their peptide bonds[46,47].

Although most of the PPIase activity in cells is found in the cytosol, family members are located in different cellular compartments. Thus cyclophilin-like proteins are also present in the periplasmic space of *E. coli* cells[48,49], and in the mitochondrial matrix of *Neurospora crassa*[50]. Furthermore, the nucleotide sequences of cyclophilin-related genes cloned from *S. cerevisiae*[51], *Drosophila melanogaster*[52,53] and vertebrate cells[54–57] and of the FKBP-related gene from human cells[58] each encode a stretch of N-terminal amino acids whose sequences are compatible with function as signal peptides for translocation into the lumen of ER. Following their translocation into the ER, some cyclophilin-related proteins may also be transported along the exocytic pathway and secreted into the extracellular medium[55,56]. Consistent with this diversity of localization are the findings that multiple genes encoding cyclophilin- and FKBP-related proteins are present in mammalian cells and lower eukaryotes[58–61], although in *N. crassa* a single gene encodes both cytosolic and mitochondrial cyclophilins[50]. Genetic studies in lower eukaryotes demonstrated that the cyclophilin and FKBP gene products that mediate sensitivity to CsA or FK506 and rapamycin are not essential for cell viability either individually[61–63] or in combination[63], either because of the presence in cells of additional proteins having PPIase activity or because this activity is not required for cell survival.

Despite our partial understanding of their adventitious role in the suppression of T-lymphocyte function[45], the real *in vivo* role and physiological substrates of enzymes with PPIase activity remain to be established. It seems likely that these highly abundant and widely distributed proteins normally act as 'conformases' (ref. 10) catalysing slow steps in the initial folding and/or rearrangement of protein structures. Initial evidence in support

TABLE 1 Enzymes and chaperones that may be involved in protein folding and assembly in cells

Organism/organelle	Enzymes			Chaperones		
Protein family	PDI	Cyclophilin PPIase	FKBP PPIase	Hsp60 (Chaperonin-60)	Hsp70 (Stress-70)	Hsp90 (Stress-90)
E. coli						
Cytosol	Thioredoxin	PPIase b		GroEL	DnaK	HtpG (C62.5)
Periplasm		PPIase a (Rotamase)				
Yeast						
Cytosol		Cph1p (Cpr1p)	Fkb1p (Fkr1p) (Rbp1p)		Ssa1-4p	Hsp83 Hsc83
ER	PDI Eug1p	yCyPB			Kar2p (BiP)	
Mitochondria				Hsp60 (Mif4p)	Ssc1p	
Drosophila						
Cytosol		CyP			Hsp68 Hsp70 Hsc1,2,4	Hsp83
ER	PDI	NinaA				
Mammals						
Cytosol		Cyclophilin (PPIase) (CyPA)	FKBP		Hsp70 (p73) Hsc70 (p72) (CUATPase) (Prp73)	Hsp90 (Hsp83) (Hsp87)
ER	PDI (ERp59) GSBP ERp72 ERp61	CyPB(rCyPLP)			BiP (Grp78)	Grp94 (ERp99) (endoplasmin)
Mitochondria				Hsp60 (Hsp58)	Hsp70 (Grp75)	
Plants						
Cytosol						
ER	PDI				b70 (BiP)	
Chloroplasts				RuSBP		

Six protein families have been identified whose members include enzymes or chaperones proposed to be involved in folding, assembly, rearrangement or degradation of proteins in cells. Members that have been characterized to date from a variety of different organisms are shown. Alternative names are shown in parentheses. For references see the text.

of this hypothesis came from the discovery that the *D. melanogaster ninaA* gene product, an eye-specific cyclophilin-related membrane protein, is required for the folding and/or stability of rhodopsins 1 and 2 (refs 52, 53, 64). Treatment of chick embryo fibroblasts with cyclosporin A delays the folding of the triple helix of type I collagen, indicating a physiological role for cyclophilin PPIase in folding in the ER (ref. 31).

The mechanism by which the two classes of PPIases catalyse rotation around specific peptide bonds also remains to be determined. Human cyclophilin and FKBP display dramatic differences in substrate specificity. Although cyclophilin has a broad specificity and does not discriminate between P_1 amino-acid residues, FKBP has a narrow specificity with a preference for hydrophobic residues at the P_1 position[65]. The reason for this preference became clear with the determination of the three-dimensional structure of FKBP (refs 47, 66), as the folding topology provides a large cavity, lined with conserved aromatic residues, that serves as the active site and drug binding pocket. Mechanistic studies (reviewed in ref. 45) suggest that both cyclophilin and FKBP catalyse the interconversion of *cis*- and *trans*-rotamers of peptide substrates by noncovalent stabilization of a twisted amide transition state.

Protein chaperones

A number of other cellular proteins, now collectively known as chaperones[67], function *in vivo* not as catalysts of secondary structure formation, but rather to recognize and stabilize partially folded intermediates during polypeptide folding, assembly and disassembly. The majority of the currently identified chaperones belong to three highly conserved protein families, whose members are widely distributed from prokaryotes to plants and mammals (Table 1). In eukaryotic cells, different family members are found in different cellular compartments and organelles. As their names imply, proteins of the hsp70, hsp90 and chaperonin (groEL/hsp60) families first came to attention because of their specific induction during the cellular response of all organisms to heat shock (reviewed in refs 68–71). Nevertheless, the majority of the family members are expressed constitutively and abundantly in the absence of any stress, and genetic studies show that many of these proteins are essential for cell viability under normal conditions of growth[68,72,73]. Many hsp family members, including those that do not respond significantly to heat shock, are induced under a variety of other stress conditions[68,74,75] whose common denominator may be the accumulation of unfolded or malfolded proteins in cells[76–79]. To reduce confusion arising from the use of the term 'hsp' to describe all family members regardless of their response to heat shock, we will use the terms 'stress-70' and 'stress-90' for members of the 70K and 90K families, respectively, a nomenclature similar to that suggested by Craig[68,80] for the 70K family of *S. cerevisiae*. The 60K (groEL/hsp60) family is currently described by the term 'chaperonin-60'[81,82].

Chaperones seem to serve many functions that stem from their ability to recognize and modulate the state of folding of polypeptides within cells (Table 2). Thus members of the stress-70 family have been implicated in the stabilization or generation of unfolded protein precursors before assembly in the cytosol[83] or translocation into organelles including the ER and mitochondria[84–86], in stabilization of newly translocated polypeptides before folding and assembly[3,23,87–90], in rearrangement of protein oligomers[72,91,92], in dissolution of protein aggregates[87,93], and in the degradation of rapidly turned-over

cytosolic proteins[94,95]. Stress-90 proteins seem to stabilize a variety of target proteins in an inactive or unassembled state[68]. Chaperonin-60 proteins bind unfolded precursors before export of secretory proteins[24,26] or assembly of protein oligomers[96] in *E. coli*, and help in folding and assembly of polypeptides translocated into chloroplasts or mitochondria in eukaryotic cells[97-99]. In addition, unrelated cellular proteins that are not members of any of these stress protein families have been implicated in protein in cells. Nucleoplasmin, for which Laskey[100,101] coined the term chaperone in 1978, binds to histones and facilitates ordered nucleosome assembly in the nucleus. In the ER, the T-cell receptor-associated protein, TRAP (ref. 102) or p28 (ref. 103), is noncovalently associated with the newly synthesized CD3 chains until they assemble with other subunits of the receptor, whereas an 88K protein transiently associates with newly synthesized major histocompatability complex (MHC) class I heavy chains[104]. In *E. coli*, SecB and trigger factor also bind unfolded secretory precursors before their export across the plasma membrane[26,105-107], PapD has been proposed to act as a chaperone in the periplasmic space to enhance the folding and assembly of components of P pili[108,109], and 'scaffolding proteins' encoded by bacteriophages promote the assembly of phage coats although they are not finally incorporated into the viral particles[110]. With the realization that polypeptide folding in the cell frequently requires the assistance of chaperones, some with a broad target specificity and some dedicated to assembly of particular macromolecules, it seems inevitable that many more such proteins will be identified in the coming months and years.

The stress-70 protein family

The role of stress-70 proteins during the heat-shock response has been studied extensively for many years (for recent reviews see refs 68-71), but only recently has their importance in normal cellular processes such as protein folding, assembly, disassembly and degradation (Fig. 1) become widely appreciated[87,95,111,112]. Stress-70 family members implicated in such processes include in *E. coli*, DnaK; in yeast the cytosolic proteins Ssa1p and Ssa2p, the ER protein Kar2p and the mitochondrial protein Ssc1p; and in mammalian cells the cytosolic proteins hsp70 (p72), hsc70 (p73, clathrin uncoating ATPase) and prp73 (peptide recognition protein 73), and the ER protein BiP (also known as grp78). Although in no case is the interaction between an individual stress-70 protein and its target polypeptide understood in molecular detail, the available evidence reveals several common features that point to a conserved mechanism for the action of these ubiquitous proteins. These features include (1) differential recognition of target polypeptides and modulation of their conformation or state of assembly, (2) involvement of ATP binding and/or hydrolysis, (3) a requirement for other heat shock proteins or cellular factors, and (4) the induction of synthesis of individual stress-70 family members by the accumulation of unfolded proteins in the appropriate cellular compartment.

***E. coli* DnaK.** DnaK, originally, defined as the product of a host gene required for bacteriophage λ DNA replication in *E. coli*, also plays fundamental roles in normal cellular physiology[72]. Mutations in the *dnaK* gene result in temperature-sensitive growth of *E. coli*, overproduction of other heat-shock proteins even at permissive temperatures, impaired synthesis of DNA and RNA and a generalized defect in proteolysis. Furthermore, the synthesis of DnaK is increased as a result of the accumulation in cells of unfolded polypeptides[113] and DnaK binds foreign eukaryotic proteins expressed in *E. coli*[114]. Finally, overproduction of DnaK aids the export of lacZ hybrid proteins across the bacterial inner membrane[115]. These observations all indicate the involvement of DnaK in modulating many protein-protein interactions *in vivo*. Detailed *in vitro* studies have now illustrated the ability of DnaK to interact with either fully assembled or unfolded polypeptide substrates. During bacteriophage replication, DnaK functions together with two other heat shock proteins, DnaJ and GrpE, to release lambda P protein from the inactive pre-primosomal replication complex[72,116,117]. DnaK also functions together with DnaJ to activate RepA initiator protein for binding to the origin of replication of plasmid P1 (ref. 118). Hydrolysis of ATP, thought to be catalysed by DnaK, is required during both reactions, and *in vitro* this ATPase activity of DnaK can be stimulated up to 50-fold in the simultaneous presence of DnaJ and GrpE (ref. 119). Finally, DnaK, which often associates with *E. coli* RNA polymerase through many purification steps, can protect the enzyme from heat inactivation *in vitro*[93]. ATP is not required for the protective effect. DnaK can also reactivate heat-inactivated RNA polymerase by dissolving aggregates formed at high temperatures but this process is absolutely dependent on the hydrolysis of ATP.

Cytosolic stress-70 proteins of eukaryotes. On heat shock of mammalian cells, both the constitutively expressed hsc70 proteins and the heat-induced hsp70 proteins migrate from the cytoplasm to the nucleus where they associate with polypeptides that form an insoluble complex at the increased temperature (reviewed in ref. 111). Subsequently the stress-70 proteins also migrate to the nucleolus and associate with partially assembled proribosomes. Presumably, at elevated temperatures nuclear proteins become partially denatured, exposing hydrophobic regions that tend to interact to form insoluble aggregates. Pelham[87,120] proposed that by binding to the exposed hydrophobic surfaces, stress-70 proteins could limit such interactions and perhaps promote disaggregation. In cell extracts, the stress-70 proteins could be released from their association with nucleolar proteins by addition of ATP, but not nonhydrolysable ATP analogues[120].

In *S. cerevisiae*, the constitutively expressed cytosolic hsc70 proteins Ssa1p and Ssa2p are required, together with a separate

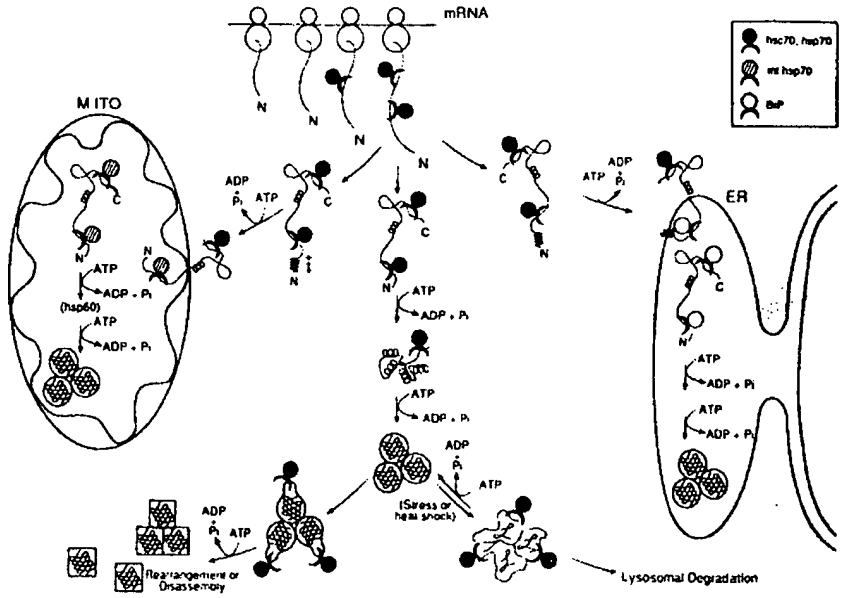


FIG. 1 Illustration of the proposed roles of stress-70 proteins in eukaryotic cells during the folding and membrane translocation of nascent polypeptides, during molecular rearrangements or disassembly, in protection from stress and in protein turnover.

TABLE 2 *In vivo* roles of protein chaperones

	Target polypeptide	Chaperone	Role	Reference
E. coli	Secretory precursors (β-lactamase, proOmpA, prePhoE)	GroEL SecB Trigger Factor DnaK?	Antifolding before translocation	24, 26, 105, 106, 107, 115, 194
	DNA replication complexes	DnaK/DnaJ	Rearrangement of protein complex	72, 116, 117
	Bacteriophage head or tail proteins	GroEL/GroES Bacteriophage scaffolding proteins	Head or tail assembly	96 110
	Foreign proteins	DnaK	Stabilization of unfolded protein?	114
	P pili in periplasmic space	PapD	Pili assembly	118
Photosynthetic bacteria	Rubisco	GroEL/GroES	Oligomer assembly	82, 188, 192, 195, 196
Chloroplasts	Rubisco	RuSBP	Oligomer assembly	81, 191, 192
Mitochondria	Mitochondrial precursors	Hsp70	Completion of translocation Stabilization of prefolded structures in matrix	90, 147 90
	Precursors in matrix	Hsp60	Stabilization of prefolded structures and folding Re-export of precursors to intermembrane space	204–206 204
ER	Nascent secretory proteins	BiP	Completion of translocation Stabilization of prefolded structures in lumen	138 3, 23, 89, 133, 134
	Mutant or foreign proteins	BiP	Stabilization of unfolded structures?	3, 151, 152
	Subunits of T cell receptor MHC Class I heavy chains	TRAP or p28 p88	Receptor assembly Stabilization of newly synthesized heavy chains?	102, 103 104
	Plant storage proteins	b70 (BiP)	Stabilization of newly synthesized polypeptides?	127, 132
Cytosol	Nascent polypeptides	Hsc70 (Hsp70?)	Stabilization of prefolded structures?	83
	Mitochondrial and secretory precursors	Hsc70 (Hsp70?)	Antifolding before translocation	84–86
	Clathrin-coated vesicles	Hsc70 (clathrin uncoating ATPase)	Binds exposed loop of clathrin light chain to promote uncoating	91, 92
	Aged? proteins	Hsc70 (Prp73)	Targeting to lysosomes for degradation	94, 95, 124
	Steroid receptors	Hsp90	Stabilizes inactive form of receptor	178–180
	Retroviral transforming proteins	Hsp90	Stabilizes inactive form of protein during transit to plasma membrane	176, 177
	Actin, tubulin	Hsp90	Stabilizes protein subunits?	68
Nucleus	Preribosomes	Hsp70/Hsc70	Protection of heat denatured proteins	111
	Histones	Nucleoplasmin	Nucleosome assembly	100, 101

N-ethylmaleimide(NEM)-sensitive cytosolic factor, for membrane translocation of secretory and mitochondrial precursors[84–86]. Consistent with this role is that the synthesis of cytosolic stress-70 proteins is induced in yeast cells by the accumulation of secretory precursors in the cytoplasm[79]. Because in cell-free experiments the need for the stress-70 proteins can be eliminated by urea-mediated unfolding of precursors[84,86], the hsc70 proteins are though to promote a translocation competent state by relaxing the tertiary structure of the polypeptides or by dissolving aggregates of untranslocated precursors[84,85]. *In vivo*, the hsc70 proteins may bind to nascent secretory precursors before they fold, maintaining them in a translocation-competent state before membrane penetration. Both Ssa1p and Ssa2p bind ATP with high affinity[84,85], but the involvement of ATP during their interaction with secretory proteins has not been established. *In vitro* experiments indicate that mammalian hsc70 can function in the same manner as Ssa1p and Ssa2p to stimulate import of M13 procoat into mammalian ER microsomes[121]. Hsc70 also interacts transiently with newly synthesized cytosolic proteins[83], presumably to facilitate their proper folding and subsequent assembly in the cytoplasm.

In vivo the targets of the cytosolic stress-70 proteins are not limited to damaged proteins or nascent polypeptides as hsc70, in the form of clathrin-uncoating ATPase[122,123] also promotes the disassembly of clathrin cages by displacing triskelions from the clathrin lattice in a process that requires ATP hydrolysis and the presence of clathrin light chains (reviewed in ref. 91). Transient changes in Ca^{2+} and/or K^+ concentrations in the cytosol surrounding a newly invaginated clathrin-coated vesicle apparently cause exposure of a stretch of amino acids (residues 47–71) in LC_a light chains that comprise a binding site for hsc70 (ref. 92). The interaction between the LC_a peptide and hsc70,

which *in vitro* both alters the conformation of the hsc70 molecule and stimulates its ATPase activity, then initiates the uncoating process which proceeds in a cooperative manner. That the hsc70 binding site on the LC_a molecule is cryptic under the ionic conditions normally present in the cytosol explains why the large cellular excess of hsc70 over clathrin does not lead to a permanent state of clathrin disassembly[92].

Finally, prp73, a stress-70 family member that is almost certainly identical to hsc70, is involved in the lysosomal degradation of intracellular proteins[94]. Prp73 binds peptide sequences (KFERQ and related sequences[95,124]) that target intracellular proteins for lysosomal degradation in response to serum withdrawal. When lysosomal uptake and degradation of protein substrates is reconstituted *in vitro*, prp73 stimulates degradation in an ATP-dependent manner[94]. Serum starvation causes induction of prp73, which presumably alters the conformation of KFERQ-containing proteins so that they can be translocated across the lysosomal membrane.

Stress-70 in the endoplasmic reticulum. The ER of eukaryotic cells contains a roughly 78K member of the stress-70 protein family[79,125-127], now named BiP[23,88]. In mammalian cells, this protein was originally described independently as the immunoglobulin heavy chain binding protein[22,128], and the glucose-regulated protein, grp78 (ref. 129). In yeast cells, BiP is the product of the *KAR2* gene[79,126], one of a class of genes involved in nuclear fusion following mating of yeast cells[130,131]. Under normal growth conditions, BiP is synthesized constitutively and abundantly and comprises about 5% of the lumenal content of the ER of mammalian cells. Its synthesis can be further induced by the accumulation of secretory precursors[79] or mutant proteins in the ER (refs 78, 132), or by a number of different stress conditions[75] that also lead to the accumulation in the ER of unfolded polypeptides[78]. BiP associates transiently with a variety of nascent wild-type exocytotic proteins[3,23,89,133,134] and more permanently with malfolded or unassembled proteins whose transport from the ER is blocked[3,23,135]. Complexes between BiP and nascent secretory proteins, isolated from extracts of mammalian cells, can be dissociated *in vitro* by the addition of ATP, but not of nonhydrolyzable analogues or ADP (ref. 125). BiP, an essential protein in yeast[79,126,136], is therefore thought to have a role in the folding and assembly of newly synthesized proteins in the ER lumen[3,23,87,88]. Initial suggestions that BiP might recognize and retain unfolded proteins in the ER[3,23,87,137], or target malfolded proteins for destruction[135] are no longer tenable. Thus, although unassembled immunoglobulin heavy chains are secreted if the BiP-binding domain is deleted[137], a truncated form of influenza haemagglutinin that does not bind BiP is neither secreted, nor degraded more slowly, than other transport-defective haemagglutinin mutants that do bind to BiP (M. Segal, J.F.S. and M.-J.G., unpublished results). In addition to modulating protein folding in the ER lumen, BiP may be directly or indirectly involved in translocation of precursors across the ER membrane. Yeast cells expressing a temperature-sensitive *kar2* mutant accumulate secretory precursors in the cytosol at the nonpermissive temperature[138]. In addition, *KAR2* interacts genetically with *SEC63* (J. Vogel and M. Rose, personal communication), a yeast gene that encodes a transmembrane protein required for ER translocation of secretory precursors[139]. Sec63p spans the ER membrane and extends into the lumen a domain containing sequences homologous to DnaJ (ref. 140), another heat-shock protein frequently required for DnaK's function in *E. coli*[72,118,119]. Whether cooperation with DnaJ-related proteins is required for the chaperone activity of all stress-70 family members remains to be determined.

Stress-70 in the mitochondrial matrix. Stress-70 proteins have been identified in the mitochondria of a number of organisms including *S. cerevisiae*[141], *Euglena gracilis*[142], *Trypanosoma cruzei*[143] and mammals[144,145]. The yeast protein (Ssc1p) has been localized to the mitochondrial matrix[90] where it performs an essential function, as disruption of the *SSC1* gene is lethal[146]. The amino-acid sequence of Ssc1p is more closely related to DnaK than are those of the other eukaryotic stress-70 proteins[141], consistent with the presumed endosymbiotic origin of mitochondria. Studies using a temperature-sensitive *ssc1* yeast mutant[90] demonstrate dual functions for Ssc1p that parallel those suggested for the ER-located BiP protein. These are involvement in translocation of precursor proteins through the lipid bilayer at mitochondrial contact sites, and involvement in folding of the imported polypeptides in the mitochondrial matrix. The translocation defect can be circumvented *in vitro* by artificially denaturing the precursor molecules, allowing investigation of events in the matrix of isolated mitochondria containing mutant forms of Ssc1p. Interestingly, the imported precursors remain in an unfolded state and can be isolated in physical association with the mutant Ssc1p protein. On the basis of these observations, Kang *et al.*[90] proposed that Ssc1p binds the precursor polypeptide as it emerges on the matrix side of the translocation apparatus in contact sites[97], supporting the continuation of translocation by 'pulling' the precursor into the matrix space. Subsequently, Ssc1p would maintain the imported precursor in an unfolded state until it is released, possibly in an ATP-dependent step, for subsequent folding catalysed by other components in the mitochondrial matrix (see Fig. 2 and discussion below of the role of the chaperonin hsp60). The isolation by crosslinking of complexes containing a partially translocated precursor, a mitochondrial outer membrane protein (ISP42) and Ssc1p provides direct evidence that Ssc1p can interact with polypeptides before the chains have been completely imported into the mitochondrial matrix[147].

Role of ATP binding and hydrolysis by stress-70 proteins. All stress-70 family members bind ATP and a number of them, including DnaK (ref. 16), hsc70 (ref. 122) and BiP (ref. 148), have weak ATPase activities that can be elicited by appropriate protein substrates and by some but not all synthetic peptides[92,148]. Adenine nucleotide binding apparently causes conformational changes in stress-70 proteins that result in altered sensitivity to proteases[149], or in alteration of their oligomeric state[150]. ATP and ADP differ in their effects. ATP protects a roughly 60K fragment of BiP while ADP protects a roughly 45K fragment; ATP stabilizes the monomeric form of hsc70 whereas ADP stabilizes the dimer. These consequences of ATP binding do not require hydrolysis as the nonhydrolysable analogue ATPγS can substitute. By contrast, addition of ATP, but not of nonhydrolysable analogues, to cell extracts causes dissolution of complexes between stress-70 proteins and their polypeptide substrates, including DnaK and bacteriophage λP protein[117], hsp70 and heat-shocked nuclei[126], hsc70 and mutant forms of the cellular p53 protein[114], and BiP and immunoglobulin heavy chains[125]. These observations led Pelham[87,120] to propose that ATP hydrolysis causes conformational changes in stress-70 proteins that are transmitted to the substrates, promoting their folding or weakening their interactions with other polypeptides. However, it is currently believed that binding of stress 70 proteins may simply stabilize unfolded conformations of their target polypeptides, preventing the formation of inappropriate intra- or intermolecular interactions. Rothman[16] has proposed that ATP hydrolysis, which takes place *in vitro* with a turnover time of about 5 min, provides a timed mechanism of release of the stress-70 protein from its substrate, freeing the polypeptide to continue the folding process. This turnover time could of course be altered *in vivo* by interactions with other cellular components. As discussed earlier, the ATPase activity of DnaK can be stimulated 50-fold in the presence of DnaJ and GrpE proteins[119].

Substrate recognition by stress-70 proteins. The molecular basis of substrate recognition by stress-70 proteins remains a matter for speculation. Prp73 binds cytosolic proteins that contain sequences identical or closely related to the consensus pentapeptide KFERQ[94,95,124]. Other stress-70 family members interact with a variety of polypeptides that do not contain any conserved

motif, suggesting that their 'recognition signals' do not consist of unique linear amino-acid sequences. Despite their broad patterns of polypeptide recognition, these proteins are not interchangeable. Thus clathrin cages elicit the ATPase activity of hsc70 but not that of BiP (ref. 148), and neither DnaK nor BiP can replace prp73 in promoting targeting of RNAase A for lysosomal degradation[94]. Evidence is lacking, however, for evolution of structural features that limit recognition by individual stress-70 proteins to *bona fide* targets, as DnaK will bind eukaryotic cellular or viral proteins when they are expressed in *E. coli*[114], and BiP will recognize bacterial enzymes or virally encoded nuclear antigens when they are artificially introduced into the ER lumen by virtue of addition of an *N*-terminal hydrophobic signal sequence[151,152]. Nevertheless, target recognition by these proteins is not indiscriminate: many examples exist of authentic secretory proteins that do not seem to associate with BiP. Similarly, high-affinity BiP binding may be confined to specific domains in individual polypeptides. Thus the CH_1 domain of immunoglobulin heavy chain is necessary for stable interaction with BiP (ref. 137), and binding of BiP to influenza haemagglutinin is apparently limited to sequences in the stem domain (M. Segal, J. S. and M.-J.G., unpublished results). In these cases BiP seems to interact with sequences that form subunit interfaces because the CH_1 domain of the immunoglobulin heavy chain is the site for docking of the light chain[137], whereas the haemagglutinin trimer assembles through cooperative folding of sequences in the stem domains[3,153]. Some proteins or protein domains may lack appropriate recognition signals and never interact with BiP. Alternatively, BiP may be involved in the folding of all nascent molecules in the ER but some interactions may be too transient to be detected experimentally, possibly because BiP interacts with different segments of polypeptides with a spectrum of affinities. In support of this latter hypothesis, BiP and hsc70 display marked differences in affinity for synthetic peptides as measured in a peptide-dependent ATPase reaction[146]. In a small set of randomly chosen peptides, a range of at least 1,000-fold of Michaelis constant (K_m) values was obtained. Unfortunately, no pattern could be discerned that correlates any sequence or structural features of the peptides with their binding affinities.

In some cases (notably that of BiP and nascent secretory proteins), stress-70 proteins discriminate between folded and unfolded polypeptides, showing no propensity to associate with native protein structures. In other cases (for example DnaK with DNA replication complexes; hsc70 with clathrin cages) these proteins interact with apparently fully folded substrates and function to alter protein-protein contacts in multisubunit complexes. These schemes are not necessarily mutually exclusive, as DnaK and hsc70 also bind to unfolded polypeptides[72,114], and BiP may interact with protein components of the ER translocation system[138]. To unify these observations, Rothman[16] has suggested that the stress-70 proteins (and the other chaperones) be regarded as polypeptide-chain binding proteins (PCBs) with peptide-binding sites that can interact with chain segments only when they are part of incompletely folded structures or when they extend as loops from otherwise fully folded proteins. This hypothesis can explain how stress-70 proteins can interact with broad specificity with unfolded polypeptides whereas individual 'folded' proteins (such as clathrin light chains[92]) can use a bait of extended chain to suborn the chaperone activity of a particular stress-70 family member to a specific purpose.

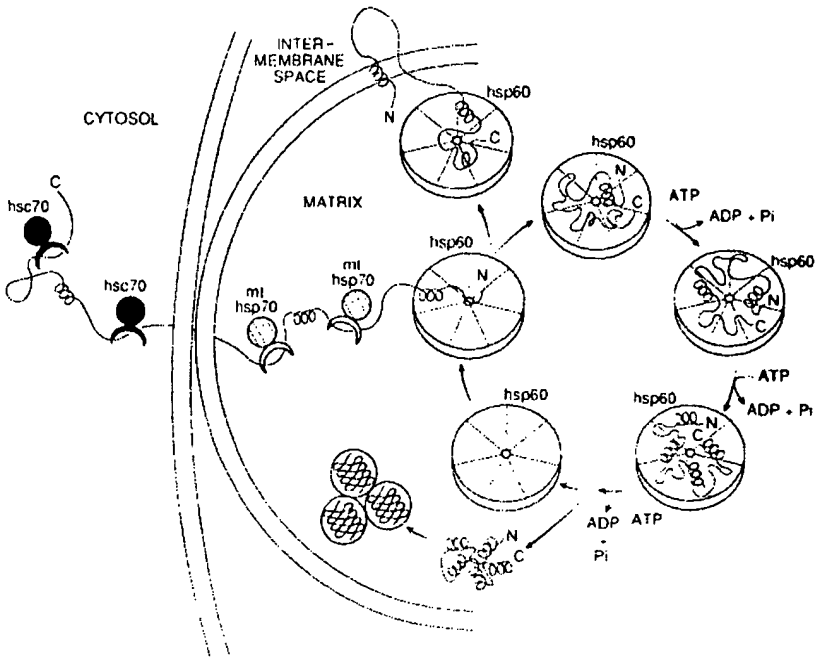


FIG. 2 Illustration of the proposed roles of hsp70 and chaperonin molecules during the import of mitochondrial precursors, their subsequent folding in the mitochondrial matrix and their reexport to the intermembrane space. The diagram is based on Fig. 1 from Neupert *et al.* (ref. 172) and in addition reflects multiple steps of folding on the chaperonin surface as suggested by Martin *et al.* (ref. 199).

Structural conservation of stress-70 proteins. Members of the stress-70 protein family have been highly conserved throughout evolution. DnaK, the single stress-70-related protein of *E. coli*, has about 50% amino-acid sequence identity with stress-70 proteins of eukaryotes[154], which are encoded by multiple *HSP70* genes that share between 50 and 95% identity at the nucleotide level[68]. Comparison of all the known amino-acid sequences of stress-70 family members reveals that the N-terminal two-thirds (about 450 amino acids) of these proteins are much more highly conserved than the C-terminal portions, suggesting a conserved domain followed by a variable region[155,156] (Fig. 3*a*). In addition, some stress-70 proteins contain short N-terminal or C-terminal extensions required for targeting to, or retention in, the appropriate cellular compartment. Thus BiP contains a cleavable N-terminal hydrophobic signal sequence specifying import into the ER[79,125] and a C-terminal tetrapeptide (for example KDEL in mammalian BiPs, HDEL in yeast BiP) that is partly responsible for retention of the protein within the ER lumen[157]. Similarly, stress-70 members located in mitochondria contain a hydrophilic N-terminal extension required for import into that organelle[141].

ATP binding and hydrolytic activity are retained by a roughly 44K, N-terminal proteolytic fragment of bovine clathrin-uncoating ATPase (hsc70), although the ATPase activity is uncoupled from its normal dependence on clathrin binding[155]. Similar N-terminal fragments are generated after proteolytic digestion of mammalian[149] and yeast[79] BiP proteins. The three-dimensional structure of the N-terminal fragment of bovine hsc70 has recently been solved to a resolution of 2.2 Å[158], revealing that the ATPase domain consists of two lobes with the nucleotide bound at the base of a deep cleft between them (Fig. 3*b*). Surprisingly, the folding topology of the hsc70 ATPase domain is nearly identical to that of the globular G actin monomer, despite there being little sequence homology between the two proteins[159]. G actin monomers contain one molecule of noncovalently bound ATP, which is hydrolysed to form bound ADP and inorganic phosphate when G actin polymerizes to

form filamentous actin. The tertiary structure of the nucleotide-binding core of the hsc70 fragment is also similar to that of hexokinase, although the remainder of the structures of the two proteins are completely dissimilar[158].

The C-terminal domain of hsc70, which has been proposed[155] to be the 'specificity' domain that couples binding of target proteins to the ATPase activity of the conserved N-terminal domain, may 'dock' onto a face of the N-terminal domain that is lined with amino-acid residues that are highly conserved between stress-70 proteins[158]. Although the structure of this portion of the molecule has not yet been determined for any stress-70 protein, two groups have now presented hypothetical models for the hsc70 C-terminal domain[160,161]. Rippman *et al.*[160] deduced a consensus secondary structure for the C-terminal domains of 33 stress-70 proteins and obtained a pattern of helices and β-strands that could be aligned with that of the α-1 and α-2 domains of the human MHC class I antigen HLA. Flajnik *et al.* found that the same domains of an MHC class I protein from *Xenopus laevis*[162] have a low amount of sequence identity with the C-terminal sequences of hsc70 and BiP proteins; they then showed that secondary structure predictions and hydropathy analyses for the corresponding regions yield very similar results if a few gaps or insertions are introduced to optimize the alignment[161]. These findings prompted both groups to model the hsc70 C-terminal domain using the known three-dimensional structure of the human class I molecule[163] (Fig. 3*c*). The putative peptide binding cleft in each of the hypothetical structures is lined with both hydrophobic and hydrophilic residues. If the peptide binding domain of stress-70 proteins does indeed closely resemble that of HLA, it is very likely that the polypeptide chain would bind in an extended conformation[164]. In fact, preliminary studies using NMR (nuclear magnetic resonance) indicate that DnaK binds a 13-residue synthetic peptide in a conformation that lacks any defined structural features[165].

A common mechanism for stress-70 action? From this abundance of disparate observations on the interaction of stress-70 proteins with their targets emerges a working model for a common mechanism for stress-70 action. Stress-70 proteins interact (probably through their C-terminal domains) with unfolded segments of polypeptide chain (Fig. 1). These unfolded segments may be presented either as nascent polypeptides emerging from the ribosome or from the lipid bilayer after membrane translocation, as sequences exposed by partial protein denaturation following an environmental stress, or as peptide loops extended from an otherwise native protein molecule. Sequence variability in the C-terminal domains of different family members may determine differences in the specificity of stress-70-peptide interactions. Although the basis for this specificity is not understood in any detail, it is likely that each stress-70 protein binds different polypeptide segments with a wide spectrum of affinities. Low-affinity binding may be reversed quickly and spontaneously, whereas release of peptide segments that are bound with high affinity may involve ATP hydrolysis mediated by the N-terminal domain of the stress-70 protein, or be effected through intervention of another cellular component (for example DnaJ, the NEM-sensitive factor, or both). Once released, the polypeptide chain has the opportunity to complete its folding by forming intramolecular interactions, or to assemble into oligomeric structures with nearby polypeptide chains, or to engage with the appropriate membrane translocation machinery or with another chaperone such as hsp60 (see below). If such interactions are not formed rapidly, stress-70 proteins, which are present at high concentration, may rebind and again stabilize the unfolded protein. During each interval of release, the polypeptide may either fold productively or form nonproductive intra- or intermolecular interactions yielding misfolded or aggregated molecules. It seems that the amino-acid sequences of wild-type proteins have evolved such that, under normal physiological conditions, productive folding or rebinding to stress-70 molecules are more likely events than misfolding. Thus wild-type polypeptides can be maintained in an assembly competent state for very long periods in the absence of their appropriate homologous or heterologous partner subunits. Similarly, cycles of binding, release and rebinding will extend the period of interaction of stress-70 proteins with polypeptides that are unable to fold productively. Sequence alterations (amino-acid substitutions, deletions or insertions), aberrant post-translational modification (for example glycosylation) or conditions of stress (high temperature, altered redox potential, decreased Ca^{2+} concentration) would pertub the normal folding pathway and increase the probability of misfolding. In such circumstances intervention by stress-70 proteins may be able to delay, but not prevent, the formation of misfolded and/or aggregated structures that are dead-ends off the folding pathway. This aberrant folding could result in occlusion of the available sites on the polypeptide for the stress-70 protein, rendering the chaperone's action less effective. Such substoichiometric binding of BiP to aggregates of nonglycosylated haemagglutinin molecules has been observed in *in vivo* experiments[135].

It is obvious that any increased probability of misfolding could be reversed by increasing the local concentration of stress-70 protein, and it is equally clear that the cell has evolved mechanisms to sense increased amounts of nascent or unfolded proteins in different cellular compartments and to respond by inducing the transcription of the appropriate stress-70 gene. For example, in *E. coli* accumulation of unfolded proteins causes increased synthesis of DnaK (and other heat-shock proteins)[113,166]. In eukaryotic cells, accumulation of unfolded proteins[76] or secretory precursors[79] in the cytosol results in induction of hsp70 and/or hsc70 proteins, whereas accumulation of unfolded proteins in the ER causes induction of BiP[78,79]. Increased concentrations of individual, constitutively expressed stress-70 proteins can be achieved by accelerating their rates of synthesis; in cases where closely related family members can perform the same or similar tasks[167], synthesis of the constitutively expressed proteins may be augmented by *de novo* induction of closely related family members. In no case do we yet understand the nature of the induction signal generated by the presence of unfolded proteins or the pathways of their transduction to the nucleus (across the ER membrane in the case of BiP).

Finally, it has been suggested that stress-70 proteins assist correct polypeptide folding not only through their 'anti-folding' function but also by disentangling malfolded or aggregated proteins using the energy released during ATP hydrolysis[87,120,168]. In cell-free translocation studies, precursors that have acquired defined structures after *in vitro* translation require unfolding for import into mitochondria[2,169-171] or ER microsomes[84,85,121]. As translocation in these systems is dependent on the presence of stress-70 proteins[84-86,121] and ATP[121,170], it was suggested that the stress-70 proteins might be supplying the unfolding activity. But it is now thought[172] that ATP hydrolysis may be required to release bound cytosolic stress-70 proteins and that the unfolding activity observed in cell-free extracts may be supplied by other components, such as the NEM-sensitive factor[86,121] or proteins present on the mitochondrial surface[172]. The only stress-70 protein for which unfolding activity has been directly demonstrated *in vitro* is DnaK which can dissolve aggregates of heat-inactivated RNA polymerase[93]. This process, which is dependent on the presence of hydrolysable ATP, requires at least stoichiometric amounts of DnaK. It is therefore possible that rather than actively unwinding the polypeptide chains in a catalytic process, DnaK might bind to peptide loops that are transiently exposed during 'breathing' of the structures of the malfolded proteins and, in a mechanism that parallels its putative role during normal folding, stabilize the polypeptide chain in a state competent for subsequent refolding to the correct conformation. ATP hydrolysis would then be required to promote release of DnaK to allow the polypeptide to continue the folding process.

The stress-90 protein family

Stress-90 proteins constitute a second major family of stress proteins (Table 1) whose members are present in all prokaryotic and eukaryotic organisms so far tested (reviewed in ref. 68). Members of the family show sequence conservation similar to that of the stress-70 family, there being greater than 40% identity between the various eukaryotic stress-90 proteins and the *E. coli* homologue HtpG. These proteins are present in high abundance under normal growth conditions, but can be further induced by heat shock or other forms of stress. They are less numerous than the stress-70 proteins, there being only one member of the family in *E. coli and D. melanogaster*, and two members in *S. cerevisiae* that differ in their constitutive level and degree of inducibility[68]. Vertebrate cells contain an additional stress-90 protein, variously name grp94 (ref. 173), ERp99 (ref. 174) or endoplasmin[175], which like BiP in the stress-70 family is synthesized as a precursor that contains an N-terminal signal sequence for ER translocation and a C-terminal KDEL tetrapeptide.

The cytosolic stress-90 proteins, whose apparent M_rs vary from 87–92K, associate with a diverse range of cellular proteins including retroviral transforming proteins, steroid hormone receptors, cellular protein kinases, actin and tubulin (reviewed in ref. 68). The common feature of these interactions seems to be the stabilization of the target proteins in an inactive or unassembled state. Thus the association of monomeric hsp90 (and an unidentified 50K phosphoprotein) with pp60^{src} immediately after its synthesis stabilizes the transforming protein in an inactive state until it reaches its appropriate destination at the plasma membrane[176,177]. Concommitant with its release from association with hsp90, pp60^{src} is phosphorylated on tyrosine, inserted into the plasma membrane, and activated as a kinase. Hsp90 is involved both in the initial folding of steroid hormone receptors[178] and in subsequent modulation of their DNA binding and transcriptional regulatory activities[179]. Thus, aporeceptors synthesized in the absence of hsp90 lack transcriptional enhancement activity and responsiveness to hormonal activation. Under normal conditions, newly synthesized aporeceptors form a complex with a dimer of hsp90 in which the receptor is stabilized in a partially unfolded conformation that is unable to bind DNA (ref. 180). Binding of steroid hormone promotes dissociation of hsp90 from the complex and allows the receptor to bind DNA. When hsp90 is removed from the complex using high salt or temperature, the interaction of the receptor with DNA is unregulated, occurring in the presence or absence of hormone. Although binding of hsp90 has been mapped to a stretch of about 80 residues in the hormone binding site on the steroid receptor[180–182], the features that determine the specificity of binding and release of target polypeptides by members of the stress-90 family are not understood. The hsp90 binds ATP in a cation-dependent manner and undergoes autophosphorylation[183]. But by contrast to stress-70 proteins and the chaperonins (see below), stress-90 proteins do not seem to catalyse the hydrolysis of ATP.

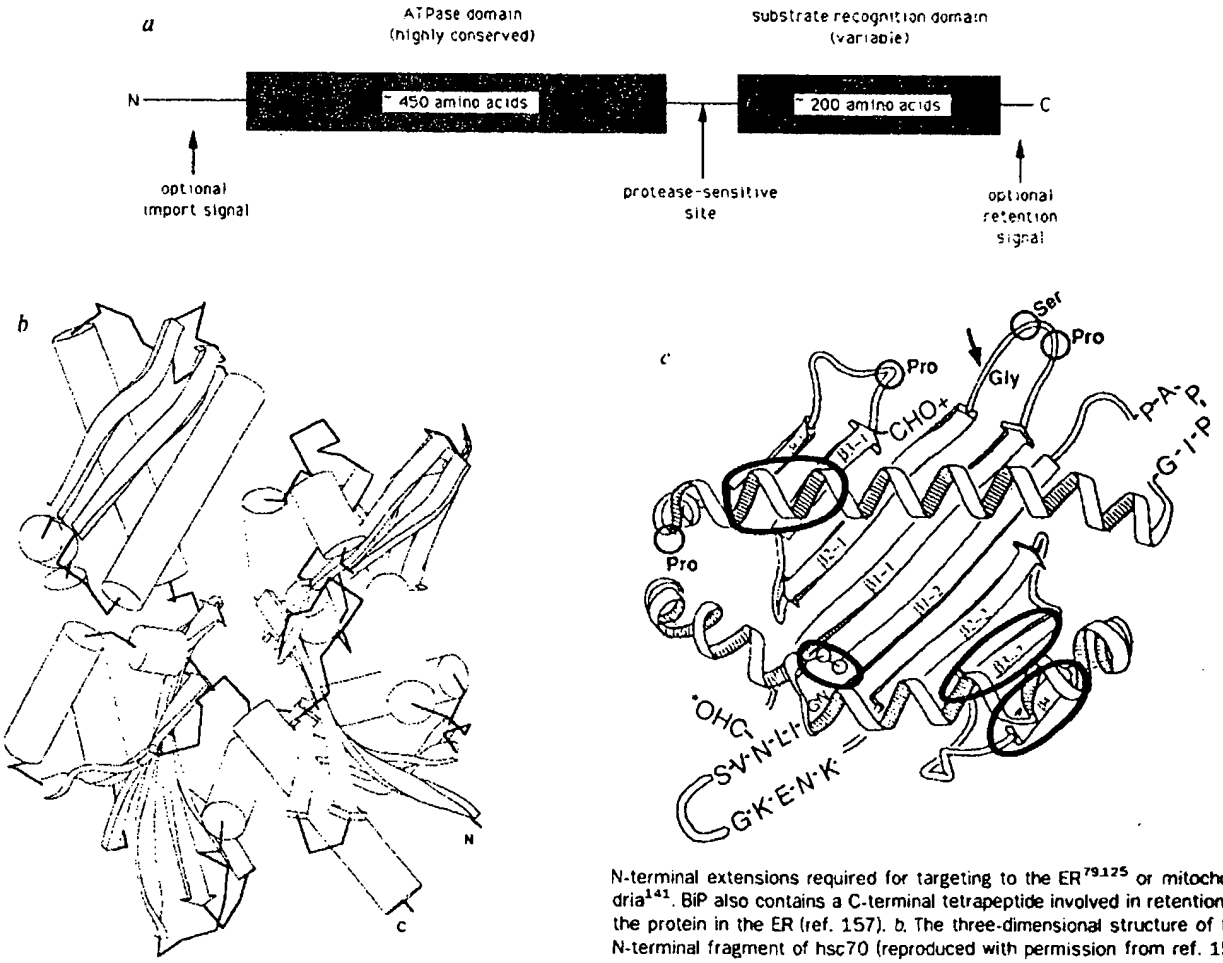


FIG. 3 Structure of stress-70 proteins. *a*, A linear diagram of a generalized stress-70 protein indicating two major domains: a highly conserved N-terminal domain that has ATPase activity and a less conserved C-terminal peptide-binding domain[155]. Some stress-70 proteins also contain short N-terminal extensions required for targeting to the ER[79,125] or mitochondria[141]. BiP also contains a C-terminal tetrapeptide involved in retention of the protein in the ER (ref. 157). *b*, The three-dimensional structure of the N-terminal fragment of hsc70 (reproduced with permission from ref. 158) reveals that the ATPase domain consists of two structural lobes with the nucleotide bound at the base of a deep cleft between them. *c*, Possible structure of the C-terminal peptide binding domain of hsc70 (reproduced with permission from ref. 161) based on the known three-dimensional structure of the human MHC class I antigen HLA (ref. 163).

Nothing is known about the function of grp94, one of the most abundant proteins resident in the ER lumen. Like BiP, grp94 is induced by the accumulation of unfolded proteins in the ER (ref. 78), suggesting that it may function with BiP to assist the assembly of nascent polypeptides. Several abundant ER-resident proteins, including grp94 and BiP, are high capacity, low-affinity Ca^{2+}-binding proteins[184], but the functional significance of this property is not understood.

The GroEL/chaperonin family

The term chaperonin was suggested by Ellis[81] to describe a class of molecular chaperones that are homologous in structure to *E. coli* GroEL. Members of this protein family are present in all prokaryotes and in those organelles of eukaryotic cells, such as mitochondria and chloroplasts, that have a probable endosymbiotic origin (Table 1). These proteins, which have been renamed chaperonin-60 (ref. 82), are large oligomers composed of 14 subunits each about 60K, arrayed as two stacked rings of seven subunits[185,186]. In *E. coli*, GroEL interacts functionally in an ATP-dependent manner with GroES (chaperonin-10), a roughly 10K polypeptide that forms a single ring of seven subunits and is the second product of the GroE operon (reviewed in ref. 96). In the absence of unfolded protein substrates, the inherent ATPase activity of GroEL is inhibited by GroES[187,188]. Mitochondria of mammalian cells contain a polypeptide that is structurally and functionally homologous to GroES (ref. 189, and P. Vütanen, unpublished results).

Remarkably, the general features of the interactions of chaperonin-60 molecules with their target polypeptides are very similar to those of stress-70 proteins, despite great differences in their sequences and oligomeric structures (for recent reviews and references see 67, 72, 96, 190–192). Thus both types of chaperones are highly abundant proteins whose rate of synthesis can be further induced by environmental stresses such as heat shock. Members of both families have been implicated in the assembly of nascent protein subunits into macromolecular structures, as well as in a number of other fundamental cellular processes. Chaperonin-60 molecules, like stress-70 proteins, bind ATP with high affinity and have weak ATPase activity and both types of proteins in some circumstances function together with other heat-shock proteins or cellular factors. Most importantly, both seem to act on their targets by stabilizing the conformation of folding intermediates, thereby preventing the formation of aberrrant structures and directing the polypeptides down biologically productive assembly pathways.

GroEL. In *E. coli*, GroEL and GroES are abundant heat-shock proteins that are also required for viability under normal growth conditions[73]. *GroE* mutants have phenotypes reminiscent of those of DnaK mutants[72]. Thus mutants lacking either GroEL or GroES have reduced rates of DNA and RNA synthesis, are blocked in cell division at nonpermissive temperatures, and show a reduction in overall protease activity. The GroE proteins are also required for bacteriophage morphogenesis in *E. coli*. Overproduction of both GroEL and GroES, but not of either alone, can suppress temperature-sensitive mutations in a large number of different genes[193] apparently by promoting the correct folding or assembly of the mutant polypeptides. Like stress-70 proteins, the GroE proteins also have a role in secretion. In bacterial cell-free protein translocation reactions, GroEL binds newly synthesized secretory precursors stabilizing them for membrane transit[24]. *In vitro*, GroEL forms complexes with unfolded precursors of several secretory proteins including β-lactamase, proOmpA and prePhoE[24,26]. *In vivo*, GroEL is required for export of β-lactamase but not other secretory precursors[194], perhaps because other prokaryotic chaperones such as trigger factor and secB function in its place[26]. Finally, overproduction of GroEL in *E. coli* can facilitate the export of lacZ hybrid proteins[115].

Insight into the mechanism of action of the *E. coli* GroE proteins has come from studies of their role in promoting the folding and/or assembly of a number of enzymes, including prokaryotic ribulose bisphosphate carboxylase (Rubisco)[82,188,195,196], pre-β-lactamase[197], citrate synthase[198], dihydrofolate reductase[196,199] and rhodanese[199,200], and of GroEL itself[201]. Partially folded protomers of these proteins form stable binary complexes with GroEL, in a process that competes both with biologically unproductive aggregation of the polypeptide chains and with their spontaneous refolding. GroEL does not interact with native proteins or with irreversibly denatured and aggregated molecules, but rather binds to labile folding intermediates likely to correspond to 'molten globules'[14,15] or 'compact intermediates'[13]. In the absence of GroES, hydrolysis of ATP by GroEL promotes the discharge of the binary complex to release partially folded but catalytically inactive polypeptides. Whether release results in the generation of native, enzymatically active molecules depends on the nature of the polypeptide substrate and is related to the propensity of each polypeptide chain to fold spontaneously under the reaction conditions employed. If GroEL is available for rebinding after release, aggregation or continued folding of the polypeptides will be inhibited or delayed[199]. In the presence of GroES, ATP hydrolysis-dependent folding occurs at the surface of GroEL through intermediate conformations that are progressively more compact but still enzymatically inactive[199]. Finally, the polypeptide is released from the complex in a form that is apparently committed to completion of folding to the native state. In every case studied the overall effect of the coordinated action of the two GroE proteins is to increase the efficiency of refolding compared with that of the spontaneous process. But how the presence of the chaperonins influences the rate of the folding reaction varies significantly from protein to protein. Thus the rate of folding of Rubisco is enhanced 10-fold relative to the spontaneous process[188], whereas the rates of folding of citrate synthase[198] and pre-β-lactamase[197] are unchanged, and those of DHFR and rhodanese[199] are decreased. These differences are likely to be related to how the specific interaction between the polypeptide chain and the chaperonin promote or interfere with rate-limiting intramolecular interactions that normally take place during the spontaneous folding process.

No more than one or two molecules of unfolded polypeptide are bound to each oligomeric assembly of 14 GroEL molecules[26,197,199]. This stoichiometry might suggest that the GroEL protomers in each heptameric ring, or in the tetradecamer, interact to form a single binding site. Alternatively, steric hindrance could limit the access of more than one or two protein molecules to 14 identical sites located in or near the hole in the centre of the doughnut-shaped structure[185,186] (Fig. 4). A polypeptide could then be bound at up to 14 sites, each of which is capable of interacting with one of the multiple recognition motifs that may be exposed on a partially folded polypeptide. This second possiblity is compatible with the suggestion that, in the presence of GroES, the folding of DHFR and rhodanese on the surface of GroEL occurs by progressive, ATP hydrolysis-dependent release of different portions of the bound polypeptide[199]. The role of GroES would then be to modulate or coordinate the ATPase activity of each GroEL protomer, perhaps to prevent all sites discharging simultaneously leading to premature release of an only partially folded molecule[199]. Overall the result of the interaction between GroEL and GroES would be to prolong contact between the chaperonin complex and its substrate as long as the polypeptide exposes structures recognized by GroEL.

Plastid Rubisco subunit binding protein (RBP). The chaperonin-60 present in chloroplasts of higher plants is a nuclear-encoded protein first identified because of its involvement in the assembly of the hetero-oligomeric plant Rubisco[81,191,192]. In contrast to GroEL and mitochondrial hsp60s, each of which contain only one type of 60K subunits, RBP consists of two distinct but related subunits, α (61K) and β (60K), probably arranged as two layers of seven monomers each[81]. The wheat Rubisco small

subunit expressed in *E. coli* associates with GroEL, providing evidence for functional homology between GroEL and wheat RBP, which have 46% identity in amino-acid sequence. As several polypeptides associate with RBP after import into isolated chloroplasts[25] it seems likely that further studies will demonstrate the involvement of RBP in the assembly of plastid macromolecules other than Rubisco.

Mitochondrial hsp60. Mitochondria from *Tetrahymena*, yeast, maize, *Xenopus* and human cells also contain proteins that are related structurally and immunologically to the GroEL protein[185,202]. A nuclear gene (*HSP60* or *MIF4*) encodes the mitochondrial homologue in *S. cerevisiae* (hsp60)[203], which displays 54% and 43% amino-acid identity, respectively, with GroEL and the α component of chloroplast RBP. Hsp60 is expressed constitutively and is localized in the mitochondrial matrix[204]. In cells expressing a mutant (*mif4*) form of hsp60, subunits of mitochondrial enzymes fail to assemble into the appropriate macromolecular complexes despite being translocated into the mitochondrial matrix and undergoing normal processing in that compartment[204]. This is the case not only for proteins normally destined for the matrix, such as the β-subunit of F_1-ATPase and hsp60 itself[205], but also for proteins with complex presequences, such as cytochrome b_2 and the Rieske Fe/S protein, whose final destination is the intermembrane space. These latter proteins accumulate as incompletely processed import intermediates[204]. Hsp60 has been directly implicated in folding of proteins in the mitochondrial matrix[206]. When assembly of precursors imported into wild-type mitochondria is arrested by depletion of ATP, by NEM treatment, or at low temperature, the unfolded polypeptides are associated with hsp60 in a high-M_r complex. Addition of ATP allows at least partial refolding of the polypeptides but does not promote their release from the complex. An unidentified factor seems to be required for the release reaction. Fig. 2 illustrates our current view of the role played by stress-70 and chaperonin molecules during the translocation and folding of polypeptides imported into mitochondria.

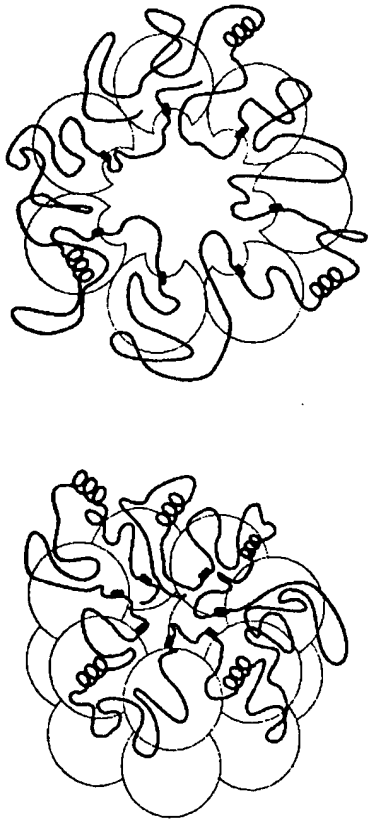


FIG. 4 Hypothetical model of the GroEL chaperonin structure reflecting multiple binding sites for a single polypeptide chain. The oligomeric structure is shown in views from the top and the side. The small corkscrews represent α-helices that may be recognition elements for binding to GroEL; the larger corkscrews represent general secondary structures in the bound polypeptide.

Mechanism of action of chaperonins? Although the broad features of the interaction of chaperonins with their polypeptide substrates have been illuminated by *in vitro* studies with GroEL and GroES (see above), many questions remain concerning the molecular details of the mechanism by which chaperonins promote protein folding and assembly. The quaternary structure of chaperonin oligomers has been revealed by electron microscopy[185,186], but nothing is known about the tertiary structures of chaperonin-60 or chaperonin-10 protomers or how they interact in the complex. Nor do we understand how chaperonins recognize their target polypeptides. The available evidence suggests that chaperonin-60 molecules bind to structural elements that are displayed by compact folding intermediates of a broad range of polypeptide substrates but absent or inaccessible in the native or aggregated forms of these proteins. The nature of these structural elements is not yet known, although preliminary experiments using NMR indicate that GroEL binds synthetic peptides that have the potential to form amphipathic α helices[207]. It will be important to learn whether chaperonin-60 molecules, like stress-70 proteins, display marked sequence specificity for binding. If so, recognition motifs on a folding protein may have evolved with a hierarchy of affinities that directs their order of release from the chaperonin complex. If so, rather than facilitating folding in a merely permissive fashion, the interaction between chaperonin and substrate may influence the pathway and the kinetics of the folding process. Finally, we need to understand the role that ATP hydrolysis plays in the folding and/or release of bound polypeptides and the manner in which chaperonin-10 regulates both this ATPase activity and the release reaction.

Roles of chaperonins and stress-70 proteins

Although stress-70 proteins and chaperonins share many common features in their modes of action, they do not perform interchangeable roles. Despite their coexistence in bacteria and in mitochondria and plastids of eukaryotic cells, each type of chaperones is independently essential for cell viability.

In mitochondria, both types of chaperones interact with unfolded or prefolded molecules: hsp70 molecules associate with imported polypeptides even before their translocation is completed[147], whereas hsp60 molecules become involved at a later stage[90,204]. It is therefore possible that stress-70 molecules interact with less folded structural elements (perhaps segments of extended polypeptide chain[16]), whereas chaperonin-60 molecules associate with structural features common only to folding intermediates[196]. Interestingly, two-dimensional NMR studies reveal that a synthetic peptide is bound by DnaK in an extended form, and by GroEL in an α-helical conformation[165].

Third, the two types of chaperones seem to play different roles during the process of polypeptide folding. Thus stress-70 molecules are thought to stabilize unfolded forms of their substrates; folding is envisaged as occurring after release from the chaperone. On the other hand, at least partial folding of polypeptides may take place on the surface of chaperonin-60 oligomers[82,199,206]. This distinction may be a consequence of the dramatic difference in the oligomeric state of the two types of molecules. Stress-70 proteins are thought to contain a single peptide-binding site and to associate as monomers with their target polypeptides[91]. Therefore, any folding step that is inhibited by chaperone binding must occur after dissociation of the complex. By contrast, the functional form of the chaperonin-60 molecule is an oligomer of 14 subunits, which apparently binds no more than 1 or 2 target polypeptides[26,197,199]. Multivalent binding of a protein to the chaperonin-60 oligomer would allow release and partial folding of one portion of the chain while the polypeptide remains attached at other sites. The extent of folding of any protein on the chaperonin-60 surface might then depend on the role in the folding process of the portion of the chain that is bound with highest affinity.

Fourth, although stress-70 proteins such as DnaK and hsc70 are known to interact both with unfolded polypeptides and with protein oligomers (such as replication complexes or clathrin triskelions), evidence is lacking for any *in vivo* role for chaperonins in the rearrangement of oligomeric complexes.

Finally, although stress-70 and chaperonin-60 molecules function alongside each other in bacteria and in the matrix of mitochondria and plastids, chaperonins have not been identified in other compartments of eukaryotic cells. Whether stress-70 family members manage alone in supporting polypeptide folding and assembly in those compartments, or whether other proteins fulfil the function of the chaperonins remains to be determined.

Future directions

We now appreciate that chaperones are involved at all stages of cellular metabolism, during protein biosynthesis and maturation, in protection from environmental stress, in rearrangements of cellular macromolecules during functional cycles of assembly and dissassembly, and finally in targeting proteins for degradation. Much progress has been made in characterizing chaperones that are members of three families of major stress proteins, and in identifying a number of unrelated proteins that also regulate or facilitate polypeptide folding in the cell. The major challenge now lies in elucidating the specific molecular mechanisms by which chaperones recognize their target proteins and promote, inhibit or reverse folding and assembly.

Although acceptance of the involvement of chaperones has required some revision of long-held views about the spontaneity of the process of protein folding, there was less resistance to a role during folding for enzymes such as PDI and the PPIases. Despite advances in our understanding of their mechanisms of action, important questions remain to be answered about the *in vivo* function and substrates of these enzymes and of a variety of newly discovered related proteins. □

Mary-Jane Gething and Joseph Sambrook are at the Department of Biochemistry, University of Texas Southwestern Medical Centre, 5323 Harry Hines Boulevard, Dallas, Texas 75235, USA. M.J.G. is an Investigator of the Howard Hughes Medical Institute.

1 Randall, L. L. & Hardy, S. J. S. *Cell* **46**, 921-928 (1986).
2 Eilers, M. & Schatz, G. *Nature* **322**, 228-232 (1986)
3 Gething, M. J., McCammon, K. & Sambrook, J. *Cell* **46**, 939-950 (1986).
4 Gething, M. J., McCammon, K. & Sambrook, J. *Meth. Cell Biol.* **32**, 185-206 (1989).
5 Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535-542 (1977).
6 Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. in *Crystallographic Data Bases: Information Content, Software Systems, Scientific Applications* (eds Allen, F. H., Bergerhoff, G. & Sievers, R.) 107-132 (Data Commission of the International Union of Crystallography, Bonn, Cambridge and Chester, 1987).
7 Anfinsen, C. B., Haber, E., Sela, M. & White, F. H. *Proc. natn Acad Sci U.S.A.* **47**, 1309-1314 (1961).
8 Anfinsen, C. B. *Science* **181**, 223-230 (1973).
9 Baldwin, R. L. *Trends biochem. Sci.* **14**, 291-294 (1989).
10 Fischer, G. & Schmid, F. X. *Biochemistry* **29**, 2205-2212 (1990).
11 Kim, P. S. & Baldwin, R. L. *A. Rev. Biochem.* **59**, 631-660 (1990).
12 Jaenicke, R. *Prog. Biophys. molec. Biol.* **49**, 117-237 (1987).
13 Creighton, T. E. *Biochem. J.* **270**, 1-16 (1990).
14 Kuwajima, K. *Proteins struct. funct. Genet.* **6**, 87-103 (1989).
15 Ptitsyn, O. B. *J. Prot. Chem.* **6**, 272-293 (1991).
16 Rothman, J. E. *Cell* **59**, 591-601 (1989).
17 Freedman, R. B. *Trends biochem. Sci.* **9**, 438-441 (1984).
18 Colman, A., Besley, J. & Valle, G. *J. molec. Biol.* **160**, 459-474 (1982).
19 Hendershot, L. M. *J. Cell Biol.* **111**, 829-837 (1990).
20 Copeland, C. S., Doms, R. W., Bolzau, E. M., Webster, R. G. & Helenius, A. *J. Cell Biol.* **103**, 1179-1191 (1986).
21 Mitraki, A. & King, J. *BioTechnology* **7**, 690-697 (1989).
22 Haas, I. G. & Wabl, M. *Nature* **306**, 387-389 (1983).
23 Bole, D. G., Hendershot, L. M. & Kearney, J. F. *J. Cell. Biol.* **102**, 1558-1566 (1986).
24 Bochkareva, E. S., Lissin, N. M. & Girshovich, A. S. *Nature* **336**, 254-257 (1988)
25 Lubben, T. H., Donaldson, E. K., Viitanen, P. V. & Gatenby, A. A. *Pt. Cell* **1**, 1223-1230 (1989).
26 Lecker, S. *et al. EMBO J.* **8**, 2703-2709 (1989).
27 Freedman, R. B., Bulleid, N. J., Hawkins, H. C. & Paver, J. L. *Biochem. Soc. Symp.* **55**, 167-192 (1989).
28 Freedman, R. B. *Cell* **57**, 1069-1072 (1989).
29 Lang, K., Schmid, F. X. & Fischer, G. *Nature* **329**, 268-270 (1987).
30 Schönbrunner, E. R. *et al. J. biol. Chem.* **266**, 3630-3635 (1991).
31 Bachinger, H. P. & Compton, L. A. *J. cell. Biochem. suppl.* **15Q**, 188 (1991).
32 Roth, R. A. & Pierce, S. B. *Biochemistry* **26**, 4179-4182 (1987).
33 Bulleid, N. J. & Freedman, R. B. *Nature* **335**, 649-651 (1988).
34 Lamantia, M. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 4453-4457 (1991).
35 Farquhar, R. *et al. Gene* (in the press)
36 Edman, J. C., Ellis, L., Blacher, R. W., Roth, R. A. & Rutter, W. J. *Nature* **317**, 267-270 (1985).
37 Holmgren, A. *A. Rev. Biochem.* **54**, 237-271 (1985).
38 Freedman, R. B., Hawkins, H. C., Murant, S. J. & Reid, L. *Biochem. Soc. Trans.* **16**, 96-99 (1988)
39 Geetha-Habib, M., Noiva, R., Kaplan, H. A. & Lennarz, W. J. *Cell* **54**, 1053-1060 (1988).
40 Bulleid, N. J. & Freedman, R. B. *EMBO J.* **9**, 3527-3532 (1990).
41 Wetterau, J. R., Combs, K., Spinner, S. N. & Joiner, B. J. *J. biol. Chem.* **265**, 9800-9807 (1990).
42 Lewis, M. J., Mazzarella, R. A. & Green, M. *J. biol. Chem.* **260**, 3050-3057 (1985).
43 Mazzarella, R. A., Srinivasan, M., Haugejorden, S. M. & Green, M. *J. biol. Chem.* **265**, 1094-1101 (1990).
44 Russel, M. *Molec. Microbiol.* **5**, 1607-1613 (1991).
45 Schreiber, S. L. *Science* **251**, 283-287 (1991).
46 Fesik, S. W. *et al. Science* **250**, 1406-1409 (1990).
47 Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. *Science* **252**, 839-842 (1991).
48 Liu, J. & Walsh, C. T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4028-4032 (1990).
49 Hayano, T., Takahashi, N., Kato, S., Maeki, N. & Suzuki, M. *Biochemistry* **30**, 3041-3048 (1991).
50 Tropschug, M. *et al. J. biol. Chem.* **263**, 14433-14440 (1988)
51 Koser, P. L., Sylvester, D., Livi, G. P. & Bergsma, D. *J. Nucleic Acids Res.* **18**, 1643 (1990).
52 Shieh, B. H., Stammes, M. A., Seavello, S., Harris, G. L. & Zuker, C. S. *Nature* **338**, 67-70 (1989).
53 Schneuwly, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 5390-5394 (1989)
54 Price, E. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 1903-1907 (1991).
55 Spik, G. *et al. J. biol. Chem.* **266**, 10735-10738 (1991).
56 Caron, P., Rothenfluh, A., McGlynn, E. & Schneider, C. *J. biol. Chem.* **266**, 10739-10742 (1991).
57 Hasel, K. W., Glass, J. R., Godbout, M. & Sutcliffe, J. G. *Molec. cell. Biol.* **11**, 3484-3491 (1991).
58 Jin, Y.-J. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 6677-6681 (1991)
59 Haendler, D., Hofer-Warbinek, R. & Hofer, E. *EMBO J.* **6**, 947-950 (1987)
60. Danielson, P. E. *et al. DNA* **7**, 261-267 (1988)
61. Tropschug, M., Barthelmess, I. B. & Neupert, W. *Nature* **342**, 953-955 (1989).
62 Wilderrecht, G., Brizuela, L., Elliston, K., Sigal, N. H. & Siekierka, J. J. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1029-1033 (1991).
63 Heitman, J., Movva, N. R., Hiestand, P. C. & Hall, M. N. *Proc. natn. Acad. Sci. U.S.A.* **88**, 1948-1952 (1991).
64 Stammes, M. A., Shieh, B.-H., Chuman, L., Harris, G. L. & Zuker, C. S. *Cell* **65**, 219-227 (1991).
65 Harrison, R. K. & Stein, R. L. *Biochemistry* **29**, 3813-3816 (1990).
66. Michnick, S. W., Rosen, M. K., Wandless, T. J., Karplus, M. & Schreiber, S. L. *Science* **252**, 836-839 (1991)
67. Ellis, R. J. & Hemmingsen, S. M. *Trends biochem. Sci.* **14**, 339-342 (1989)
68 Craig, E. A. *A. Rev. Genet.* **22**, 631-677 (1988)
69 Craig, E. A. *CRC Crit. Rev. Biochem.* **18**, 239-280 (1985).
70 Lindquist, S. *A. rev. Biochem.* **55**, 1151-1191 (1986)
71. Craig, E. A. & Gross, C. A. *Trends biochem. Sci.* **16**, 135-140 (1991).
72 Georgopoulos, C., Ang, D., Liberek, K. & Zylicz, M. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R. I., Tissières, A. & Georgopoulos, C.) 191-221 (Cold Spring Harbor, New York, 1990).
73 Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. *J. Bact.* **171**, 1379-1385 (1989
74 Subjeck, J. R. & Shyy, T.-T. *Am. J. Physiol.* **250**, C1-C17 (1986).
75. Lee, A. S. *Trends biochem. Sci.* **12**, 20-23 (1987).
76 Ananthan, J., Goldberg, A. L. & Voellmy, R. *Science* **232**, 522-524 (1986)
77. Hiromi, Y. & Hotta, Y. *EMBO J.* **4**, 1681-1687 (1985).
78. Kozutsumi, Y., Segal, M., Normington, K., Gething, M. J. & Sambrook, J. *Nature* **332**, 462-464 (1988)
79 Normington, K., Kohno, K., Kozutsumi, Y., Gething, M. J. & Sambrook, J. *Cell* **57**, 1223-1236 (1989).
80 Werner-Washburne, M., Stone, D. E. & Craig, E. A. *Molec. cell Biol.* **7**, 2568-2577 (1987).
81. Hemmingsen, S. M. *et al. Nature* **333**, 330-334 (1988).
82. Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884-889 (1989).
83 Beckmann, R. P., Mizzen, L. A. & Welch, W. J. *Science* **248**, 850-854 (1990).
84. Chirico, W. J., Waters, M. G. & Blobel, G. *Nature* **332**, 805-809 (1988).
85. Deshaies, R. J., Koch, B. D., Werner-Washburne, M., Craig, E. A. & Schekman, R. *Nature* **332**, 800-805 (1988).
86 Murakami, H., Pain, D. & Blobel, G. *J. Cell Biol.* **107**, 2051-2057 (1988).
87. Pelham, H. R. B. *Cell* **46**, 959-961 (1986)
88. Hendershot, L. M., Bole, D. & Kearney, J. F. *Immun. Today* **8**, 111-114 (1987).
89 Ng, D. T. W., Randall, R. E. & Lamb, R. A. *J. Cell Biol.* **109**, 3273-3289 (1989).
90 Kang, P. J. *et al. Nature* **348**, 137-143 (1990).
91 Rothman, J. E. & Schmid, S. L. *Cell* **46**, 5-9 (1986).
92 DeLuca-Flaherty, C., McKay, D. B., Parham, P. & Hill, B. L. *Cell* **62**, 875-887 (1990).
93 Skowyra, D., Georgopoulos, C. & Zylicz, M. *Cell* **62**, 939-944 (1990)
94 Chiang, H.-L., Terlecky, S. R., Plant, C. P. & Dice, J. F. *Science* **246**, 382-385 (1989).
95 Dice, J. F. *Trends biochem. Sci.* **376**, 305-309 (1990)
96 Georgopoulos, C. & Ang, D. *Semin. Cell. Biol.* **1**, 19-25 (1990).
97. Hartl, F. U. & Neupert, W. *Science* **247**, 930-938 (1990).
98 Hallberg, R. L. *Semin. Cell Biol.* **1**, 37-45 (1990).
99 Hemmingsen, S. M. *Semin. Cell Biol.* **1**, 47-54 (1990).
100 Laskey, R. A., Honda, B. M. & Finch, J. T. *Nature* **275**, 416-420 (1978).
101. Dingwall, C. & Laskey, R. A. *Semin. Cell Biol.* **1**, 11-17 (1990).
102 Bonifacino, J. S. *et al. J. biol. Chem.* **263**, 8965-8971 (1988).
103. Petley, C. L., Alarcon, B., Malin, R., Weinberg, K. & Terhorst, C. *J. biol. Chem.* **262**, 4854-4859 (1987)
104 Degen, E. & Williams, D. B. *J. Cell Biol.* **112**, 1099-1115 (1991).
105. Collier, D. N., Bankaitis, V. A., Weiss, J. B. & Bassford, P. J. Jr. *Cell* **53**, 273-283 (1988).
106 Crooke, E., Guthrie, B., Lecker, S., Lill, R. & Wickner, W. *Cell* **54**, 1003-1011 (1988).
107. Kumamoto, C. A. *Molec. Microbiol.* **5**, 19-22 (1991).
108. Holmgren, A. & Branden, C.-I. *Nature* **342**, 248-251 (1989)
109. Sambrook, J. F. & Gething, M. J. *Nature* **342**, 224-225 (1989).
110 Prevelige, P., Thomas, D. & King, J. *J. molec. biol.* **202**, 743-757 (1988)
111 Pelham, H. R. B. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R. I., Tissières, A. & Georgopoulos, C.) 287-299 (Cold Spring Harbor, New York, 1990).
112 Morimoto, R. I., Tissières, A. & Georgopoulos, C. in *Stress Proteins in Biology and Medicine* 1-36 (Cold Spring Harbor Laboratory, Cold Springer Harbor, New York, 1990).
113 Parsell, D. A. & Sauer, R. T. *Genes Dev.* **3**, 1226-1232 (1989)
114. Clarke, C. F. *et al. Molec. cell. Biol.* **3**, 1206-1215 (1988)
115 Phillips, G. J. & Silhavy, T. J. *Nature* **344**, 882-884 (1990).
116 Zylicz, M., LeBowitz, J. H., McMacken, R. & Georgopoulos, C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6431-6435 (1983).
117. Liberek, K., Georgopoulos, C. & Zylicz, M. *Proc. natn. Acad. Sci. U.S.A.* **18**, 6632-6636 (1988).
118. Wickner, S., Hoskins, J. & McKenney, K. *Nature* **350**, 165-167 (1991).
119 Liberek, K., Marszalek, J., Ang, D., Georgopoulos, C. & Zylicz, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2874-2878 (1991).
120 Lewis, M. J. & Pelham, H. R. B. *EMBO J.* **4**, 3137-3143 (1985).
121 Zimmermann, R., Sagstetter, M., Lewis, M. J. & Pelham, H. R. B. *EMBO J.* **7**, 2875-2880 (1988).
122 Chappell, T. G. *et al. Cell* **45**, 3-13 (1986)
123. Schlossmann, D. M., Schmid, S. L., Braell, W. A. & Rothman, J. E. *J. Cell. Biol.* **99**, 723-733 (1984).
124 Dice, J. F. & Chiang, H.-L. *Biochem. Soc. Symp.* **55**, 45-55 (1989).

125. Munro, S. & Pelham, H. R. B. *Cell* **46**, 291–300 (1986).
126. Rose, M. D., Misra, L. M. & Vogel, J. P. *Cell* **57**, 1211–1221 (1989).
127. Fontes, E. B. P. *et al. Pl. Cell* (in the press).
128. Morrison, S. L. & Scharff, M. D. *J. Immun.* **114**, 655–659 (1975).
129. Pouysségur, J., Shu, R. P. C. & Pastan, I. *Cell* **11**, 941–947 (1977).
130. Conde, J. & Fink, G. R. *Proc. natn. Acad. Sci. USA* **73**, 3651–3655 (1976).
131. Polaina, J. & Conde, J. *Molec. Gen. Genet.* **186**, 253–258 (1982).
132. Boston, R. S., Fontes, E. B. P., Shank, B. B. & Wrobel, R. L. *Pl. Cell* (in the press).
133. Dorner, A. J., Bole, D. G. & Kaufman, R. J. *J. Cell Biol.* **105**, 2665–2674 (1987).
134. Blount, P. & Merlie, J. P. *J. Cell. Biol.* **113**, 1125–1132 (1991).
135. Hurtley, S. M., Bole, D. G., Hoover-Litty, H., Helenius, A. & Copeland, C. S. *J. Cell Biol.* **108**, 2117–2126 (1989).
136. Nicholson, R. C., Williams, D. B. & Moran, L. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1159–1163 (1990).
137. Hendershot, L. M., Bole, D., Köhler, G. & Kearney, J. F. *J. Cell Biol.* **104**, 761–767 (1987).
138. Vogel, J. P., Misra, L. M. & Rose, M. D. *J. Cell Biol.* **110**, 1885–1895 (1990).
139. Rothblatt, J. A., Deshaies, R. J., Sanders, S. L., Daum, G. & Schekman, R. *J. Cell Biol.* **109**, 2641–2652 (1989).
140. Sadler, I. *et al. J. Cell. Biol.* **109**, 2665–2675 (1989).
141. Craig, E. A. *et al. Molec. cell. Biol.* **9**, 3000–3008 (1989).
142. Amir-Shapira, D., Leustek, T., Dalie, B., Weissbach, H. & Brot, N. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1749–1752 (1990).
143. Engman, D. M., Kirchoff, L. V. & Donelson, J. E. *Molec. cell. Biol.* **9**, 5163–5168 (1989).
144. Leustek, T., Dalie, B., Amir-Shapira, D., Brot, N. & Weissbach, H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7805–7808 (1989).
145. Mizzen, L. A., Chang, C., Garrels, J. I. & Welch, W. J. *J. biol. Chem.* **264**, 20664–20675 (1989).
146. Craig, E. A., Kramer, J. & Kosic-Smithers, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4156–4160 (1987).
147. Scherer, P. E., Krieg, U. C., Hwang, S. T., Vestweber, D. & Schatz, G. *EMBO J.* **9**, 4315–4322 (1990).
148. Flynn, G. C., Chappell, T. G. & Rothman, J. E. *Science* **245**, 385–390 (1989).
149. Kassenbrock, C. K. & Kelly, R. B. *EMBO J.* **8**, 1461–1467 (1989).
150. Schmid, S. L., Braell, W. A. & Rothman, J. E. *J. biol. Chem.* **260**, 10057–10062 (1985).
151. Sharma, S., Rodgers, L., Brandsma, J., Gething, M. J. & Sambrook, J. *EMBO J.* **4**, 1479–1489 (1985).
152. Bird, P., Gething, M. J. & Sambrook, J. *J. Cell Biol.* **105**, 2905–2914 (1987).
153. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
154. Bardwell, J. C. A. & Craig, E. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 848–852 (1984).
155. Chappell, T. G., Konforti, B. B., Schmid, S. L. & Rothman, J. E. *J. biol. Chem.* **262**, 746–751 (1987).
156. Millarski, K. L. & Morimoto, R. I. *J. Cell Biol.* **109**, 1947–1962 (1989).
157. Pelham, H. R. B. *A. Rev. Cell Biol.* **5**, 1–23 (1989).
158. Flaherty, K. M., DeLuca-Flaherty, C. & McKay, D. B. *Nature* **346**, 623–628 (1990).
159. Flaherty, K. M., McKay, D. B., Kabsch, W. & Holmes, K. C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5041–5045 (1991).
160. Rippmann, F., Taylor, W. R., Rothbard, J. B. & Green, N. M. *EMBO J.* **10**, 1053–1059 (1991).
161. Flajnik, M. F., Canel, C., Kramer, J. & Kasahara, M. *Immunogenetics* **33**, 295–300 (1991).
162. Flajnik, M. F., Canel, C., Kramer, J. & Kasahara, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 537–541 (1991).
163. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
164. Madden, D. R., Gorga, J. C., Strominger, J. F. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
165. Landry, S. J. & Gierasch, L. M. *Peptides* (in the press).
166. Goff, S. A. & Goldberg, A. L. *Cell* **41**, 587–595 (1985).
167. Craig, E. A. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R. I., Tissières, A. & Georgopoulos, C.) 320–321 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1990).
168. Rothman, J. E. & Kornberg, R. D. *Nature* **322**, 209–210 (1986).
169. Eilers, M., Hwang, S. & Schatz, G. *EMBO J.* **7**, 1139–1145 (1988).
170. Eilers, M. & Schatz, G. *Cell* **52**, 481–483 (1988).
171. Chen, W.-J. & Douglas, M. G. *J. biol. Chem.* **262**, 15605–15609 (1987).
172. Neupert, W., Hartl, F. U., Craig, E. A. & Pfanner, N. *Cell* **63**, 447–450 (1990).
173. Lee, A. S., Bell, J. & Ting, J. *J. biol. Chem.* **259**, 4616–4621 (1984).
174. Mazzarella, R. A. & Green, M. *J. biol. Chem.* **262**, 8875–8883 (1987).
175. Koch, G. L. E., Smith, M., Macer, D. R. J., Webster, P. & Mortara, R. *J. Cell Sci.* **88**, 217–232 (1986).
176. Brugge, J. S., Erikson, E. & Erikson, R. *Cell* **25**, 363–372 (1981).
177. Brugge, J. S., Yuonemoto, W. & Darrow, D. *Molec. cell. Biol.* **4**, 2697–2704 (1983).
178. Picard, D. *et al. Nature* **348**, 166–168 (1990).
179. Picard, D., Salser, S. J. & Yamamoto, K. R. *Cell* **54**, 1073–1080 (1988).
180. Dalman, F. C. *et al. J. biol. Chem.* **264**, 19815–19821 (1989).
181. Howard, K. J., Holley, S. J., Yamamoto, K. R. & Distelhorst, C. W. *J. biol. Chem.* **265**, 11928–11935 (1990).
182. Dalman, F. C., Scherrer, L. C., Taylor, L. P., Akil, H. & Pratt, W. B. *J. biol. Chem.* **266**, 3482–3490 (1991).
183. Csermely, P. & Kahn, C. R. *J. biol. Chem.* **266**, 4943–4950 (1991).
184. Macer, D. R. J. & Koch, G. L. E. *J. Cell. Sci.* **91**, 61–70 (1988).
185. McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **8**, 371–380 (1988).
186. Hendrix, R. W. *J. molec. Biol.* **129**, 375–392 (1979).
187. Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R. & Georgopoulos, C. *J. biol. Chem.* **261**, 12414–12419 (1986).
188. Vitanen, P. V. *et al. Biochemistry* **29**, 5665–5671 (1990).
189. Lubben, T. H., Gatenby, A. A., Donaldson, G. K., Lorimer, G. H. & Viitanen, P. V. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7683–7687 (1990).
190. Ellis, R. J., van der Vies, S. M. & Hemmingsen, S. M. *Biochem. Soc. Symp.* **55**, 145–153 (1989).
191. Ellis, J. *Science* **250**, 954–959 (1990).
192. Gatenby, A. A. & Ellis, J. R. *A. Rev. Cell Biol.* **6**, 125–149 (1990).
193. Van Dyk, T. K., Gatenby, A. A. & LaRossa, R. A. *Nature* **342**, 451–453 (1989).
194. Kusukawa, N., Yura, T., Chaharu, U., Akiyama, Y. & Ito, K. *EMBO J.* **8**, 3517–3521 (1987).
195. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. *Nature* **337**, 44–47 (1989).
196. Lorimer, G. H., Viitanen, P. V., van der Vies, S. M., Donaldson, G. & Gatenby, A. A. *J. cell. Biochem. suppl.* **15G**, 160 (1991).
197. Lamintet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Plückthun, A. *EMBO J.* **9**, 2315–2319 (1990).
198. Buchner, J. *et al. Biochemistry* **30**, 1586–1591 (1991).
199. Martin, J. *et al. Nature* **352**, 36–42 (1991).
200. Mendoza, J. A., Rogers, E., Lorimer, G. H. & Horowitz, P. M. *J. biol. Chem.* **266**, 13044–13049 (1991).
201. Lissin, N. M., Venyaminov, S. Y. & Girshovich, A. S. *Nature* **348**, 339–342 (1991).
202. McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **7**, 4414–4423 (1987).
203. Reading, D. S., Hallberg, R. L. & Myers, A. M. *Nature* **337**, 655–659 (1989).
204. Cheng, M. Y. *et al. Nature* **337**, 620–625 (1989).
205. Cheng, M. Y., Hartl, F.-U. & Horwich, A. L. *Nature* **348**, 455–458 (1990).
206. Ostermann, J., Horwich, A. L., Neupert, W. & Hartl, F.-U. *Nature* **341**, 125–130 (1989).
207. Landry, S. J. & Gierasch, L. M. *Biochemistry* **30**, 7359–7362 (1991).

ARTICLES

Ridges, hotspots and their interaction as observed in seismic velocity maps

Yu-Shen Zhang* & Toshiro Tanimoto

Seismological Laboratory, 252-21, California Institute of Technology, Pasadena, California 91125, USA

A new global S-wave velocity model reveals that although mid-ocean ridges and hotspots are both underlain by low-velocity anomalies in the mantle, these have distinctly different structures. This implies that there are differences between the upwelling mechanisms under ridges and under hotspots. The velocity model also shows that there may be interactions between ridges and hotspots near Afar and St Helena.

RIDGES and hotspots are two main forms of upwelling from the interior of the Earth and are essential features of global tectonics. In ocean basins, they are the dominant modes of igneous activity, and are known to produce magmas of differing, although sometimes overlapping, geochemistry[1]. The mechanisms operating below these features have been inferred from surface observations, including geochemical, topographic and geoid data, and theoretical models based on gross structure and evolution of the mantle[2–5]. Seismic data provide critical information for understanding these features from maps of three-dimensional (3D) structure. On a local scale, seismic techniques have been successful, for example, in constraining the size of magma chambers under ridges[6], but such studies are restricted to particular areas where dense observations have been made.

We have obtained a new global, 3D S-wave velocity model for the upper mantle by analysis of travel times of long-period Love and Rayleigh waves (75–250 s) using ~18,000 seismograms. It is now becoming possible for global seismic studies to detect seismic velocity anomalies caused by various tectonic features (as opposed to purely long-wavelength variations), and to make comparisons between regions in order to understand the underlying mechanisms. Details of the data analysis and S-wave velocity structure are reported elsewhere[7,8]. Here we focus on S-wave velocity variation under ridges and hotspots.

Long-wavelength features of the new model are similar to previous global, 3D seismic models[9–14]. But previous models were represented by spherical harmonic expansion only up to angular degrees 6–10, with horizontal resolution lengths of 4,000–6,000 km. The chief advance of our model, which is expanded up to degree 36, is a considerable improvement in the

* Present address: Institute of Tectonics, University of California, Santa Cruz, California 95064, USA.